REPORT NO. SSD 3166R
DCOCS-4

Prepared for:
NATIONAL AERONAUTICS AND
SPACE ADMINISTRATION
MANNED SPACECRAFT CENTER
HOUSTON 1, TEXAS

77p.

N64-16770

CODE-1

(NASA CR-55812;

STUDY ON OPTICAL COMMUNICATIONS FROM DEEP SPACE

INTERIM PROGRESS REPORT

27 March 1963 ~~through~~ 31 May 1963

2 OTS

Kenneth L. Brinkman [1963]
77p refs

OTS PRICE

XEROX $ 7.60 ph
MICROFILM $ 2.51 mf.

AEROSPACE GROUP

HUGHES

1161272

HUGHES AIRCRAFT COMPANY
CULVER CITY, CALIFORNIA

SSD-3340R)

(NASA Contract ~~No.~~ NAS 9-879)
P.R. No. 3-440-20351
WAS 10-81-02-36

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

INTRODUCTION

This Interim Progress Report presents a discussion of work performance on Contract NAS 9-879 during the period from 27 March 1963 to 31 May 1963.

Since the Final Report has been rescheduled for submission on August 15, 1963, no attempt has been made to summarize work performance in previous reporting periods as originally planned for this report. Additional study results covered in the extended period have not been delineated but are implicit in work performed on the preliminary design analysis presented here.

This analysis is performed for the DSV to Earth link using simple quantum detection, since this is the detection process most completely understood, and since the effects of atmospheric turbulence on coherent (heterodyne) reception cast serious doubts on its utility for earth-based receiver systems. These effects of atmospheric turbulence on the selection of system design parameters are discussed early in this report, as are the considerations for site selection to avoid the problem of interfering weather and cloud cover.

To serve as an illustrative vehicle for design analysis, and since these techniques appear to give the most promise at this juncture of the study, two different systems were chosen for analysis. These consist of a pulsed laser using PPM and a continuous wave laser using PCM with polarization modulation.

The pulsed laser chosen to describe was an electroluminescent diode laser since it appears to have many of the attributes required for the PPM type of operation, especially the high pulse repetition frequency and pulse-control capability, but it should be emphasized that no claim of a strong analytical basis for this choice is made. The inherent noise characteristics of this device will have to be studied further. The spatial coherence of diode lasers and their resultant ability to be collimated to the degree indicated here, as well as the power, repetition frequency, operating frequency, and efficiency assumed, must be subjected to further examination.

The CW operation described herein is also not without limitation, especially in the amount of power required. A valid comparison of the two systems depends strongly on the choice of an equivalent set of system parameters, and since in attempting to provide this basis, a certain amount of speculation had to be resorted to, this comparison should be taken as a means of illustrating the respective design difficulties, rather than as a definitive endorsement of a particular system. The systems described clearly illustrate the strong bias toward PPM systems in a low noise environment; as the noise increases, the PCM system very likely takes precedence. The system noise thus plays a significant role in the optimum system selection. The sources of noise and its effect on system performance for various physical systems using different types of modulation and coding must be examined in greater detail. This must be performed for coherent as well as incoherent (quantum counters) systems.

In this manner, additional steps toward the establishment of optimal deep space communication system can be taken so that in the next report phase, further narrowing of these differences can be established.

for R. L. Stroup

Kenneth L. Brinkman
~~Project Manager~~
NASA Deep Space Optical
Communications Study

1.0 SYSTEM DESIGN EXAMPLE

The current state of technology is such that the direct spacecraft-to-earth link will be more economical than the links using intermediate stations on an earth satellite or on the moon. Even though such stations potentially have capability for very high performance because they are not limited by the atmosphere, it is difficult at present to decide if they will have any practical (economic) advantage. Probably large manned scientific satellites and moon bases will be constructed, and then deep-space communication could readily be a secondary use for telescopes needed at such locations for various scientific studies. Since there is little data for use in solving design problems of large, multiple-use telescopes in such systems, and since the problems of most critical importance are those relating to the deep space vehicle itself, the present design example is restricted to the direct DSV-earth link.

The system to be described is based on various reasonable extensions of present technology; the system does not depend for its development on the validity of all these extrapolations, since alternatives present themselves at all points. The purpose is to convey an impression of the performance and operating characteristics of a system which could be developed to fly by approximately 1970.

1.1 SYSTEM CONFIGURATION

The general description of the system given in this section will be followed by a more detailed discussion of various factors considered in the design, and finally by a detailed discussion of the design example.

The basic task assigned the system for the present discussion is transmission of a standard television picture. Voice and telemetry channels could be encoded with the picture by using the dead-time at the end of each scan line. Since standard television sync circuits need not be employed, no interference with the picture transmission need occur even if most of the dead-time is given over to such transmission; however, probably only several kcps of signal bandwidth is needed for such voice and telemetry, and one sample per television scan line would then be adequate at a standard rate of 15,750 lines per frame.

Figure 1 shows the general form of the system. The laser transmitter in the deep-space vehicle is modulated by the television signal and sync signals to provide the signal detected by the receiver at the earth station, which must then decode the transmission. The earth station transmits a voice channel up to the deep-space vehicle for giving instructions and also to act as a tracking beacon. Each station must track the other so that the highly directional transmitted beams may be sure of reaching the intended receiver. The functions of each station thus include input, output, signal processing, transmitting, receiving, and tracking.

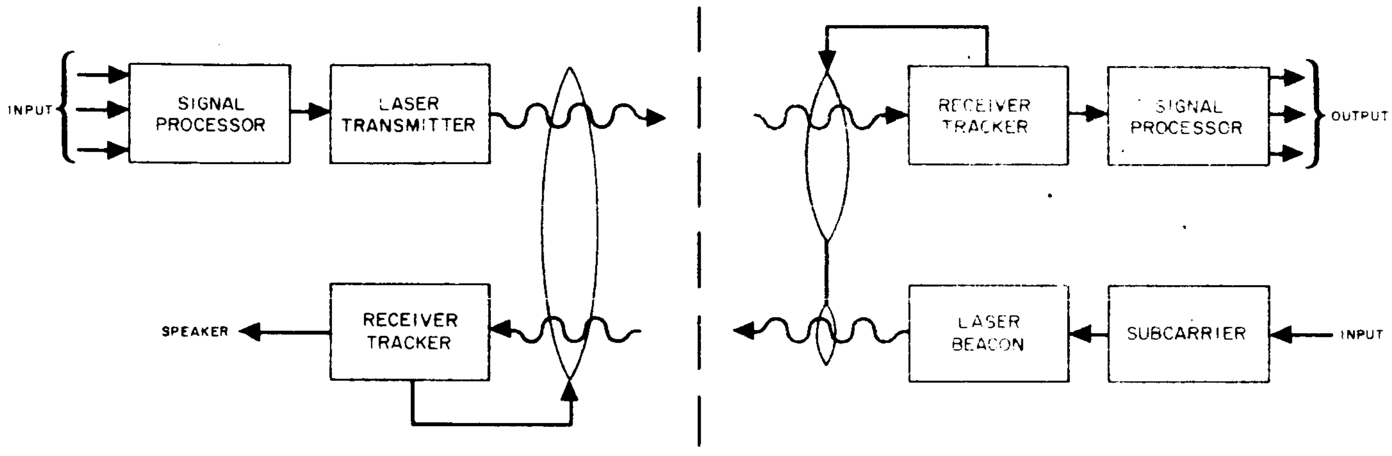
1.2 EARTH-BASED STATIONS

1.2.1 Location

Two or three airborne stations taking turns flying high-altitude circumpolar courses could maintain continuous communication without interference from cloud cover. However, the cost of constructing and maintaining such systems would be prohibitive. It is more economical to construct several ground-based stations.

The Handbook of Geophysics[1] reproduces a Weather Bureau map (Figure 2) giving the annual hours of sunshine over the world. Except for afternoon thundershowers and night fogs, which are correlated with the position of the sun, any other object in the solar system would be observable for an equal period of time each year. However, it is undesirable to work near the horizon with a precision optical system. The data have therefore been adapted, conservatively assuming the sun to be equally likely to be obscured by clouds at any time when it is above the horizon, so that they reflect the probability the spacecraft should be observable when it is above 20° from the horizon. Latitude effects were considered in calculating the spacecraft longitudes for which it would be above 20° for each site, assuming the ΔSV

1. "Handbood of Geophysics for Air Force Designers", Geophysics Research Directorate, Air Force Cambridge Research Center, Air Research and Development Command, United States Air Force, First Edition, 1957.



a) DSV station b) Earth station

Figure 1. General System Interrelations

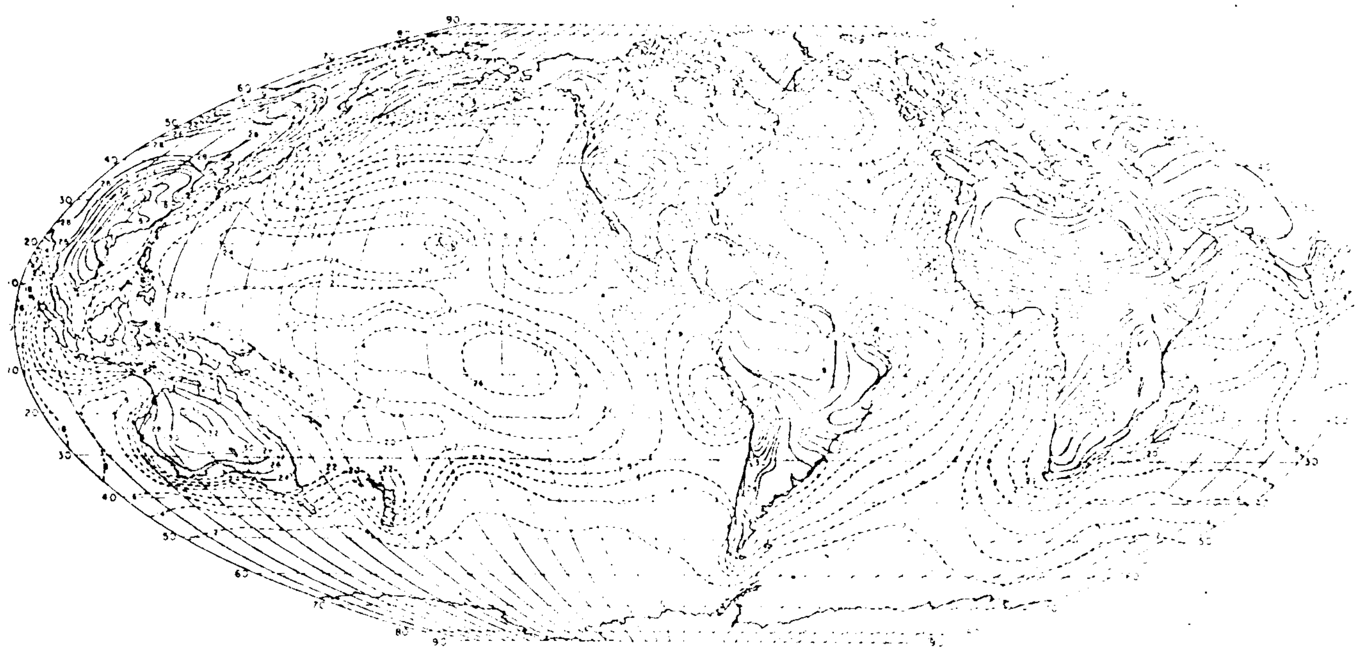


Figure 2. World Map Showing Hundreds of Hours of Sunshine per Year

is at the celestial equator. Data for eight sites are presented in Table I and Figure 3. Site No. 1 (Hawaii) may be better than indicated, because the lee side of the island of Hawaii contains a very localized desert region, which may not be represented at its best on a world map.

TABLE I.

Site No.	Longitude	Latitude	Annual Hrs. of Sunshine	Probability of Clear View	Longitude Coverage
1	19° N	156° W	2800	0.64	± 69°
2	32° N	116° W	4000	0.92	± 66°
3	23° S	69° W	3800	0.87	± 68°
4	28° S	16° E	3600	0.82	± 68°
5	21° S	25° E	4000	0.92	± 69°
6	31° N	68° E	3400	0.78	± 66°
7	44° N	116° E	2800	0.64	± 62°
8	20° S	113° E	3200	0.73	± 69°

The lowest over-all probability of system operation due to the weather occurs when only sites seven and eight are available. The probability is then 0.903. However, this occurs only for four degrees of longitude, and only if the arbitrary 20° elevation limit is held firm. The next lowest occurs when site one is also used; the probability is thereby raised to 0.965. The average for all longitudes is 0.986. This is what one might call the reliability of weather for the system.

1.2.2 Receiver Optical Instrumentation

Each station on the earth would contain a large astronomical telescope as its primary equipment. This telescope would be housed in an observatory design to run cool during the day to prevent the occurrence of rising air currents from sunlit walls; hence the major source of image disturbances in daytime operation would be eliminated. Just how effective this can be will not be known until the large solar telescope at Kitt Peak is placed in operation.

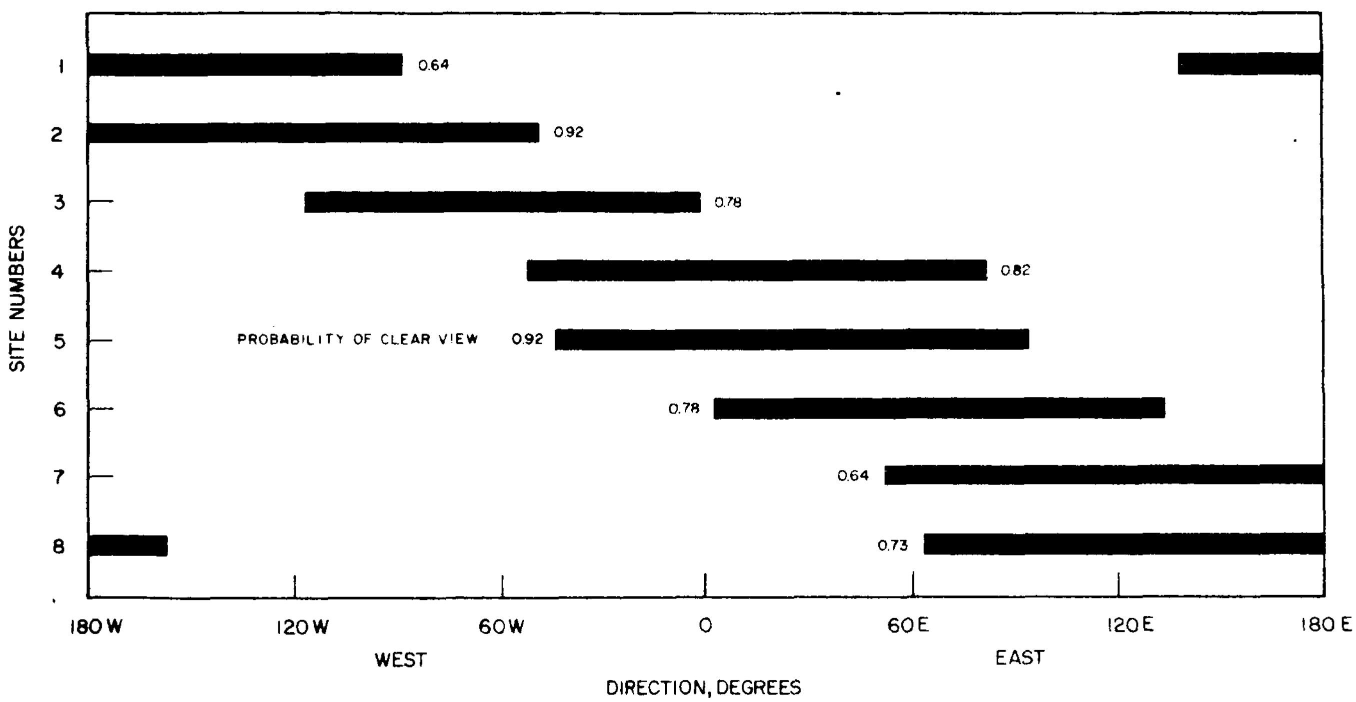


Figure 3. Longitude Coverage and Probability of Cloud-Free Operation for Eight Selected Sites

Like that telescope and other solar telescopes, this telescope would need to be located high above the ground level to reduce disturbances produced by turbulent, heated air rising from the ground. The optical quality of air during the daytime varies approximately[2] as the cube root of the height above the ground; as air rises from the ground it is mixed and becomes more homogeneous.

Even if it were necessary to reduce cost by accepting less than astronomical quality in the telescopes, they still could be used for many secondary or subsequent scientific studies. Their construction should make them desirable for use as solar telescopes, although they may have to be modified to dissipate the solar energy internally. Since they will be located in desert regions, atmospheric clarity for infrared observations should be good. Stellar and planetary observations throughout the infrared spectrum are hardly begun. Variable stars and other unique objects may be much better understood when they are studied over a larger spectral range. Much other scientific work could be done, and other communication channels handled, such as data from monitor satellites (e.g., a synchronous meteorological satellite) or moon bases. These large telescopes therefore represent more than merely the earth end of a single communication link.

An important limitation of the earth-based system is that heterodyne reception is impractical for large receiver telescope apertures because of random phase differences among the light wave fronts in various parts of the telescope aperture due to atmospheric turbulence. Heterodyne reception depends upon phase coherence between the local oscillator and the signal, and it would be quite impractical to compensate for a multitude of different phases across the aperture. Unlike a simple power detector, the heterodyne system would convert a steady signal into a much weaker and noisier signal because the voltage due to various portions of the wave front would add and substract randomly. Even for uniform wave fronts heterodyne reception is unproven as a low-signal-level technique because both theoretical and practical possibilities of generating the necessary constant-amplitude local oscillator signals are unknown. The main reason heterodyne reception is desirable is that it should

2. V. I. Tatarski, "Wave Propagation in a Turbulent Medium", translated by R. A. Silverman (McGraw-Hill Book Company, Inc., New York, 1961).

permit narrow-band, photon-noise-limited operation with solid-state detectors. These detectors have a higher quantum efficiency than photoemissive surfaces, and hence the required signal power might be substantially reduced, especially for infrared wavelengths. To offset this deficiency of the phototubes, operation in the visible region of the spectrum is presumed. At a wavelength of 6300 Å, the quantum efficiency of the tri-alkali photosurface is 0.05 and that of a silicon photodiode is approximately 0.5. Of course, direct use of the silicon diode is prohibited by its internal noise, which fortunately in heterodyning does not mix with the local oscillator frequency to appear amplified in the output because it is negligible at the required optical frequency. Laser technology is expected to extend further into the visible region, where photosurface quantum efficiency increases up to 0.3; already the Raman effect and mixing in nonlinear dielectrics have made the whole region, even into the ultraviolet, available for high-intensity pulsed operation[3].

1.2.3 Effects of Atmospheric Turbulence

Atmospheric turbulence is troublesome since it causes not only angular blurring and quivering but also fluctuations in the signal level (as in stellar scintillation) and hence causes the signals to fade. In this section, the magnitude of such fading is estimated for various conditions.

At an earth-based station, there is a fundamental difference between the fluctuations in signal level during transmission and those during reception. During reception, although diffraction at the spacecraft spreads the beam over a large portion of the earth, all the energy incident on the receiver aperture can be detected if a sufficiently large field stop (detector) is used. However, during transmission, only that portion of the beam that leaves the atmosphere in the direction of the spacecraft is used. There is only a slight spreading of the beam during passage through the atmosphere, but angular or phase disturbances are created because the plane wave front has been distorted. These disturbances will result in a large spreading of the beam after subsequent propagation. Angular divergence here, perhaps not yet affecting beam diameter because of the large initial diameter, will ultimately be the determining factor

---

3. P. M. Maker, "Nonlinear Optical Phenomena", presented May 17, 1963 in UCLA Engineering Extension short course Quantum Electronics; to be published in book form by McGraw-Hill Book Company, Inc., edited by M. L. Stitch.

in beamspread. In antenna phraseology, the top of the atmosphere is still in the near-field region of the transmitter. A large telescopic collector could be located above the atmosphere and arranged to detect the energy passing through an extremely small field stop. This large collector would be equivalent to the spacecraft, for the small field stop represents the small aperture of the spacecraft. The basic difference between transmission and reception can be summarized as follows: In transmission, cumulative phase fluctuations (which cause angular divergences) are important; however, in reception, only the cumulative amplitude fluctuations (produced by phase fluctuations near the top of the atmosphere) are significant. Although the fluctuations during reception are much smaller than those during transmission, they will be calculated first because the appropriate theory is more complete.

The study carried out by Tatarski[2] yields a useful phenomenological equation [Tatarski Eqs. (13.32) and (13.54)] that is valid for values of aperature diameter D, much greater than 10 $(\sec\theta)^{1/2}$ cm:

$$\overline{\left[\log_e(P/P_o)\right]^2} = a^2 D^{-7/3} \sec^3\theta \ ,$$

where P is the flux incident on a telescope aperture having a diameter D at a zenith angle θ; P_o is the logarithmic time average of P, i.e., $\log_e P_o = \overline{\log_e P}$; and a^2 is an empirical constant. The value of a^2 may be determined empirically. A log-normal distribution is assumed, even though it is known that for large values of D, the superposition of distributions from uncorrelated areas of the aperture tends to produce a Gaussian distribution. For small fluctuations, the difference is unimportant, and when much superposition occurs, the fluctuations tend to become small.

Tatarski (p. 237) cites scintillation data taken at the Perkins Observatory[4]. The value of

$$\overline{\left[\log_e(P/P_o)\right]^2}$$

4. W.M. Protheroe, "Preliminary Report on Stellar Scintillation", Contrib. Perkins Observ. Ser. 2, No. 4, 127 (1955).

In transmission, the first question of importance is the degree to which atmospheric turbulence spreads a beam of energy. Data taken during the site selection program for the Kitt Peak National Observatory[5] indicate that 70 percent of the light from a star falls within a 10-μrad field stop for 95 percent of the nighttime seeing conditions at this location. Extrapolation to daytime conditions is almost impossible because the difference in the amount of the incident light is primarily caused by turbulence in the air near the telescope, which depends on the construction techniques used, etc. However, not more than a tenfold increase in beam diameter due to blur can be expected if the observatory is based on a careful thermal design. This consideration also establishes the field-stop diameter used for the receiver, and hence the daytime sky radiation included with the signal

A blurred stellar image exhibits a shifting structure that is caused by interference between rays that enter through different portions of the aperture. If a laser that can generate a plane wave is used, the transmitter can project a similar image out to the spacecraft. At great distances, the receiver of the spacecraft will be small compared with the size of the structure of this projection. Just as the average contrast in Fresnel diffraction is less for large disks, the contrast in these images (and hence percent of modulation at the spacecraft) is least when the seeing conditions are worst; hence, the percent of modulation is large only when the beam is narrow and its intensity is great. Fading of the signal is therefore never as serious as would be predicted from an analysis of conditions representing simultaneously the lowest signal and highest modulation.

Unfortunately, the data needed to accurately predict the modulation due to this cause are not available. However, in view of the ease of supplying power and the comparatively low rate of data transmission, to prevent fading caused by atmospheric turbulence it should be simple to transmit signals that are ten times stronger than would otherwise be needed.

1.2.4 Transmitter Optical Instrumentation

The transmitter optics will be smaller than those of the receiver.

5. A. B. Meinel, "Astronomical Seeing and Observatory Site Selection", in Telescopes, ed. by G. P. Kuiper and B. M. Middlehurst (University of Chicago Press, Chicago, 1960).

was found to be 0.0096 in winter and 0.0064 in summer for an aperture having a diameter of 12.5 inches and presumably for small values of θ.

Hence, we take

$$a^2 = 0.0096\,(12.5)^{7/3}$$

$$a^2 = 3.5 \text{ inch}^{7/3}.$$

Therefore,

$$\overline{\left[\log_e(P/P_o)\right]^2} = 3.5D^{-7/3}\sec^3\theta,$$

where D is expressed in inches.

These data were obtained at night. Daytime conditions are similar to those at night for scintillations observed with large-aperture telescopes, although image motion and blur are much worse during the day. Although no data are available, it is probably safe to assume that a^2 is equal to 10 for daytime conditions.

For a daytime observation at a zenith angle of 60^o with a 120-inch telescope, we have

$$\overline{\left[\log_e(P/P_o)\right]^2} = 10(100)^{-7/3}(2)^3$$

$$= 1.72 \times 10^{-3}$$

$$\left[\log_e(P/P_o)\right]_{rms} = \pm\, 0.0337$$

$$(P/P_o)_{rms} \cong 1.03,\ 0.97$$

$$(P/P_o)_{3\sigma} \cong 1.11,\ 0.90\ .$$

Therefore, the scintillation is not serious for such a receiver; if the signal is made 1.11 times stronger, portions of a message will seldom be missed on account of fading. Since the fluctuations occur at low frequencies (less than 1 kc), the transmission of a short-duration pulses is unaffected by the atmospheric modulation of steady signals, such as sunlight reflected from the spacecraft, except in that the shot noise due to this background is not constant.

A six-inch aperture would be adequate to collimate most laser outputs to much better than the 0.1 mrad beam diameter expected for severe (daytime) atmospheric turbulence. This better collimation is worth while for use in good atmospheric conditions, but more important is the fact that over a large aperture the image motion effects average out to a simple blur. Another averaging effect can be obtained by using multiple apertures. The transmitter optical instrumentation could well consist of one laser divided among several collimators. Perhaps four six-inch refracting telescopes mounted in a square configuration about the tube of the comparatively large receiver telescope would be the best arrangement.

1.2.5 Function of the Station

Each earth-based station is to be capable of independently establishing communication with the DSV, given recent ephemeris data for the vehicle. After the vehicle is a few million miles from the earth its angular position against the stars can be accurately predicted days in advance. As it approaches another planet its orbit will be disturbed, but even then the trajectory can be accurately determined because observations from the vehicle itself are available to aid in determining the orbit. Consequently, no difficulty is expected in pointing at the invisible DSV to within a few seconds of arc. This is quite adequate for pointing a transmitter with a beamwidth of 0.1 mrad.

One mode of operation would be for the receiver in the DSV to be simply tracking the earth, perhaps scanning over the surface in a systematic manner. It could then lock onto the beacon formed by the transmitter on the earth, and the DSV could then begin transmitting. Meanwhile, voice commands would come up the beacon channel from the earth, and the tracker at the earth station may be locked onto the DSV transmitter. The earth station is given responsibility for initiating communication because the DSV cannot readily determine which earth station is in the most favorable position, considering cloudiness, time of day, atmospheric turbulence, etc. If several earth beacons were presented, their differences in intensity might be due to slight differences in alignment and/or collimation, and hence such differences would not be significant. However, measurements of reference stars performed at the earth stations could readily provide a basis for determining which of the stations would be best to use at the moment. Every factor except scintillation of the beacon itself could be evaluated in this manner; such scintillation would only be approximated by observations of fluctuations in the center of magnified stellar images.

The earth station would provide the data for boresighting and focusing the DSV transmitter and tracker-receiver combination, including lead-angle correction for the aberration of light. To provide the data for this, the DSV would wobble the transmitter beam about the tracker axis in a predetermined rosette or spiral pattern, and the earth station would record the received signal. Atmospheric scintillation would interfere with this operation, but if the wobble scan were sufficiently redundant scintillation could be averaged out and the true beam condition could be determined. As a result of the analysis of this data the earth station would be able to direct the DSV to make any necessary boresight adjustments, and also report on the beam diameter to verify focus or indicate focus error. This procedure would be necessary only infrequently. If the lead angle were computed from orbital data, it would not be necessary to do this more than once during the free flight of the DSV, although the verification which it provides might make it desirable as a weekly routine.

The conditions of reception of the signal must be optimized by adjustment of focus and field stop diameter so that the signal is maximized while the background radiation noise (daytime sky or planetary background) is minimized, thereby providing the most faithful and clear reception. Electronic gains and nonlinear operations on the signal must be controlled. Much of this could be done automatically, but the necessary criteria are not yet established. The type of picture to be transmitted and the purpose for which it is needed will determine some of the important specifications of the system. Many adjustments will probably be made manually while viewing the picture. It is the responsibility of the earth station not only to optimize the signal for recording and/or retransmission to other earth stations but also to notify the DSV of any deficiencies in the transmission, such as excessive clipping or failure to use the full dynamic range of the channel, irregularities in frequency control or synchronism, or irregularities in the telemetered data itself. Although all these functions could be carried out in the DSV, the earth stations can reduce the work load on the crew by thus relieving them of most such monitoring functions.

The function of an earth station is thus seen to require a complex of optical and electronic technology, although little beyond what is already regularly accomplished by various devices separately is needed.

1.3 DEEP-SPACE VEHICLE STATION

The station at the DSV end of the link is similar to the earth station in that it consists of input and output devices, a laser transmitter, a receiver, a tracker, and signal processing electronics. Two important differences are the higher data rate assumed to be required for transmission and the space environment of the system. As a result of the former a narrower beam is required, and as a result of the latter, i.e., absence of an atmosphere, such a narrower beam is possible. The details of making the use of a very narrow transmitter beam (about 0.005 mrad dia.) feasible are the major tasks of the optical and mechanical design of the spacecraft telescope. In this section the general features of the design and alternatives are discussed.

1.3.1 Optical Design

In order to ensure boresight between the transmitter, receiver, and tracker they should utilize the same primary optical system. It will be assumed that the transmitter and receiver will utilize different wavelengths but that both are near 6300 Å. Therefore the necessary separation of transmitted and received signals can only partially be accomplished spectrally. One portion of the aperture of the optical system will be used for the transmitter and another portion for the receiver and tracker, which can readily be combined together. If widely different wavelengths were used for the sight and the receiver-tracker combination the full aperture could be used for each, but then one would possibly have to operate in a less desirable spectral region. Moreover, if the full primary optical system aperture is used for the transmitter it should be free of any central obstruction, and hence probably should be refractive rather than reflective; but a refractive system might have different focus and boresight for significantly different wavelengths. A circular logic is then used to argue for aperture-sharing of reflective optics and approximately but not exactly equal wavelengths for the two directions of transmission. It should be clear that no firm requirement is established for this type of system, and further study might indicate a different preference even without introduction of new information such as laser wavelengths and collimation requirements.

It can be shown on general grounds that a laser with a diffraction-limited beam can be recollimated to produce a diffraction-limited beam of the diameter of the collimator, and hence that the beamwidth is inversely proportional to the diameter of the collimator or collimating telescope output beam. This result is even more general in that a beam which is broader than the diffraction limit by any factor will retain this factor in recollimation. Moreover, if a beam does not have equal width as measured in two orthogonal planes of symmetry containing the axis of the beam, applying cylindrical lens collimators to these two planes separately can make the beamwidth equal in the two planes. If in each plane the beamwidth approached equally close (in proportion) to the respective diffraction limit, the collimator apertures would be equal in both planes. It is expected that, in spite of the great disparity in beamwidth in the two planes, for the electroluminescent diode laser the aperture difference required to collimate to equal beamwidths will not be very great. If the beam is diffraction-limited in one plane and has a beamwidth 1.6 times the diffraction limit in the other plane, a rectangular aperture 5 in. x 8 in. would suffice for collimation to 0.005 mrad in both planes.

Because the field of view of the three devices (transmitter, receiver, and tracker) is quite small, a section of the internal beam a few inches from focus will approximate an image of the aperture. A small diagonal mirror, subtending only a part of this aperture, may be used to introduce the transmitter. The remainder of the aperture is then available for the receiver-tracker combination.

1.3.2 Tracking Techniques

Many techniques are available for generation of the required tracking signals. A restriction of this system which is not applicable in general to tracking devices is that the beacon modulation carrying the voice channel should not be disturbed. This could be accomplished by simply dividing the signal into two beams, but it is also possible to develop tracking techniques which do not spoil the signal channel. One such approach is to divide the channel into two parts in such a way that the difference or ratio between them represents the tracking error by its amplitude (or duration) and phase (or time delay) and then to detect both parts so that the sum will still accurately represent the transmitted signal.

Image motion for sensing of tracking error by an a-c technique could be introduced by a magnetic, electrostrictive, or magnetostrictive drive applied to a lens supported on a compliant suspension. Therefore no bearings are needed, and there are no "moving parts" in the ordinary sense. High reliability can therefore be achieved. The amplitude of the motion required is small, and such a technique is entirely practical. A somewhat more difficult tracking problem was solved by a conceptually similar approach for the star tracker in Hughes Aircraft Company's preliminary proposal on the Orbiting Astronomical Observatory, and working models of this device were made.

The selection of the optical design is somewhat influenced by the nature of the mounting on the spacecraft. For example, an internal mounting in which the system looks out through a window would be possible if the same face of the vehicle were always kept facing earth. In that case a refractive objective might be preferred because this would keep the mass closer to the window and, if balance is desired because of small vehicle accelerations or shifts in vehicle attitude or center of mass, the gimbal center could be kept close to the window, and hence the required window size would be minimized. However, it is more likely that large angular coverage will be required, and also the window is undesirable because of its susceptibility to erosion and damage by micrometeorites. Refractive optics used externally would be subject to the same erosion. Although some shielding could be provided for refractive optics, reflective optics could be somewhat more readily protected because the main mirror normally lies at the bottom of the telescope tube. Moreover, (for a manned spacecraft) if they are arranged for convenient access, the aluminized or silvered surface could be chemically cleaned and, using the space vacuum to accomplish the necessary evacuation, could be quite simply renewed by evaporation. Naturally pits in the glass would be permanent; after aluminizing such pits would scatter light which they had formerly merely absorbed, but much damage consists of very minute pock marks which would not penetrate a moderately thick aluminum layer and hence could be removed by this technique. Only very long-duration flights should experience such difficulties anyway.

1.3.3 Solar Effects

For a mission to Venus stations on the earth would at times have to look through the solar aureole, where the sky brightness is up to 1000 times that of the average over the sky, because at closest approach Venus is approximately between the earth and the sun. Probably a mission would be at the planet of interest at the time of closest approach. The system will not have to look directly at the solar disk, for the planet will not make another transit until the year 2004, and will not come within a degree of the limb of the sun until 1988. In any case the planet will not spend more than six days within five degrees of the sun, so that if communication is restricted to simple telemetry and voice during this period the loss of data would not be extensive.

For a mission to Mars this situation is rather different, and creates thermal and mechanical design problems for the spacecraft telescope system. Consider a mirror at the bottom of a tube of height equal to twice its diameter. When the earth is about 15 degrees from the sun the latter will illuminate one side of the telescope tube externally and the other side internally. This internal illumination will probably fall on blackened surfaces, and hence will cause substantial heating. A portion of the main mirror will also be illuminated. Such partial illumination will cause thermal expansion of that portion of the mirror with respect to the rest, and hence will cause distortion of the mirror, even if it is made of fused silica. Thermal differences will be exaggerated by the insulation provided by the space vacuum. It is thus more difficult to look near the sun than directly at the sun; in the latter case the mirror is uniformly heated, and only a shift of focus occurs. Another problem is the presence of the reflected beam; this forms a concentrated image of the sun, and can damage portions of the telescope, such as the supports of the secondary mirror, if allowed to strike them. Moreover, scattered light in the telescope, although tolerable, introduces some noise photons into the receiver and tracker.

A solution of these problems might be accomplished by carefully designing the system to withstand such distrubances. However, a more satisfactory solution might be to provide a structure to shade the telescope from the sun. It might be that some portion of the spacecraft, such as a solar cell panel or other solar energy unit could be utilized. It would be possible to provide a telescoping sunshade which would form a long extension of the main telescope tube. Most effective of all would be a plate mounted on the end of a long rod, forming a sort of parasol.

1.3.4 Mechanical Design

Without specification of the dynamic environment of the spacecraft, i.e., angular orientation, rates, and accelerations and linear accelerations, it is difficult to select a tracking servo technique and hence determine requirements on the mechanical design. If the spacecraft is large and is stably oriented it might be possible to operate a simple position servo directly referred to spacecraft coordinates. In the more general case a two- or three-axis isolation gimbal would be needed to prevent motions of the spacecraft from disturbing the telescope of the communication system. Such a gimbal would provide a platform aligned (within some tracking error) with the line-of-sight to the earth. The telescope would then only need to make small corrections in order to be accurately pointed. It would be decoupled from the platform position so that sudden disturbances in platform orientation would not disturb pointing accuracy. Other tracking approaches are possible, such as having, instead of the separate isolation gimbal, an optical correction made by motion of a relay lens in the optical system. This would lead to a quick-response servo, which hardly seems appropriate for tracking such steady objects as the earth beacons. The mechanical design thus depends upon many factors not yet specified. Many requirements for compensation of deflections and damping of vibrations remain to be specified, and hence the finesse required in the design of an actual system can hardly be guessed at now, and the alternatives are too numerous to discuss.

2.0 INFORMATION THEORY CONSIDERATIONS

The value of information theory as a tool in analysis of communication systems rests in its ability to describe in a quantitative way the accuracy with which the received data represents the intended message. The concept of a bit, representing the designation of 1 of 2 equally likely possibilities, is the accepted quantitative measuring unit for information. The requirement for the rate at which the communication channel should be capable of transmitting information could then be specified as a capacity in bits/second. This method of measuring the quality of transmission can be applied with some confidence to the transmission of quantitative data. For example, a telemetry channel capable of sending Y numbers per second with 3 digit accuracy should have an information capacity of 10Y bits/second (since $\log_2 (1000) \cong 10$). In the application of this method of analysis to pictorial or television transmission, it is admittedly less descriptive of the ability of the channel to reproduce the important elements of the picture. The usual method, which is used here, is to divide the picture into resolution elements and examine the average degree of determination of the correct intensity on the individual elements of the transmitted picture. In the terminology of information theory this degree of determination is expressed in bits/symbol, a quantity which will be denoted by H. A rough correspondence between gray scale and the quantity H may be noted. If H bits per symbol are transmitted then the intensity is equivalently specified by designating one of 2^H quantized levels to represent it. Since, in the process of coding the video waveform for transmission, the amplitude of the signal may be transformed by nonlinear signal processing into intensity gradations corresponding to the recognition levels of the viewer, the quantity 2^H corresponds approximately to the effective number of gray scale levels present in the transmitted picture.

The average information rate is given in terms of the bandwidth and H by

$$R = 2\,B_r\,H \quad \text{bits/second.}$$

Estimates of real time information rates and bit/symbol requirements were presented in an earlier report[6] for several different types of data. The

6. Interim Progress Report for the period 7 Nov. 1962 to 7 Jan. 1963, pp. 22-25.

limiting information rate of a communication channel (channel capacity) with an average power limitation in the presence of additive random noise is given by the Shannon limit $C = B_r \log_2 (1 + SNR)$, where the signal-to-noise power ratio is given by the average square signal amplitude to mean square noise amplitude ratio in the receiver. The signal amplitude for the optical communication system is proportional to the radiant power of the transmitted beam which is proportional to the current (voltage) produced in the detector. In order to avoid ambiguity in the use of the term power, (since power in the detector is proportional to radiant power squared in the beam) it is convenient to speak in terms of the nondimensional units of signal quanta. This will also be necessary for the technical reason that the number of quanta available to transmit each datum point from maximum distance is small enough that the discrete nature of the signal becomes important. In relation to an optical channel using, for example, a multiplier phototube detector, signal quanta refer to the photoelectrons emitted from the photocathode. The rate of flow of signal quanta, $\dot{S}$, may then be referred to the optical beam power received. The relation is

$$\dot{S} = \eta \, T_o \, A \, J \quad \lambda/hc$$

where

η = detector quantum efficiency

T_o = transmission of optical system

A = collector area

J = signal irradiance (watts/cm^2)

λ/hc = number of photons per joule of incident radiation

The relation to the anode current, i, is

$$\dot{S} = i/Ge$$

G = multiplier current amplification

e = charge on an electron

The noise in an optical channel is due to statistical fluctuations in the signal as well as the rms fluctuations in external noise sources (usually shot noise produced by background radiation or thermionic photocathode dark current). If a constant signal is transmitted and the average received signal quanta rate $\dot{S}$ is sampled for periods of length τ a large number of times the measurements will give a Poisson distribution about the mean $S = \dot{S}.\tau$. The rms amplitude deviation is equal to $\sqrt{S}$ which may be regarded as a signal noise. If the average number of noise quanta received during the sampling period is N then, since the signal and noise are independent distributions, the total rms noise amplitude is $\sqrt{S + N}$. Thus, for an optical channel, the effective noise competing with the signal increases with the instantaneous signal level. If an amplitude modulated optical system the instantaneous signal to noise ratio is $S^2/(S + N)$. This causes the uncertainty in the measurement of large-amplitude signals to be much worse than for small-amplitude signals. When the required number of distinguishable levels in the amplitude becomes large the signal must be increased in proportion to the square of number of levels, and hence becomes very large even in the absence of external noise sources. From the standpoint of efficiency this consideration strongly favors transmitting each symbol with pulses of small amplitude rather than one pulse of large amplitude.

In the following sections, the information rates of specific optical communication channels are computed. The quantities of interest are the number of signal quanta necessary to achieve a required number of bits/symbol and/or to provide an allowable error rate. In general, the types of laser modulation which are practical for a deep space communication channel limit the peak signal power available for the transmission of a symbol. The average transmitter power requirement for the system is proportional to the product of the peak power per symbol and the rate at which the symbols are transmitted.

2.1 INFORMATION RATES OF PULSE POSITION MODULATED OPTICAL CHANNELS

In pulse position modulation (PPM) the signal waveform is sampled at equidistant points as shown in Figure 4. The signal is then coded to give transmission of short pulses of standard height whose position in the sampling time interval carries the information about the height of the sampled waveform. There are several pulse coding techniques available and the selection of a code depends on trade-offs between bandwidth, peak signal to noise, and average transmitter power requirements. Perhaps the most straightforward coding method consists of sending a single pulse whose position (i.e., time delay) measured from the leading edge of the interval is proportional to the height of the signal waveform. The gray scale which can be transmitted is limited by the length of the sampling interval and the time resolution.

The sampling interval may be regarded as consisting of K distinguishable positions. The number K is determined by the communication bandwidth and sampling interval τ. For transmission of television data K corresponds to the maximum available gray scale resolution. When the communication link operates ideally, the number of bits of information transmitted per sampling period (bits/symbol) is $H_o = \log_2 K$. In the operation of the channel the information transmitted is less than H_o due to presence of noise and statistical fluctuations in the signal itself. When the receiver threshold is established at a level V_B, there is a probability T_S that the threshold will be exceeded at the position of the signal and a probability T_N that the threshold will be exceeded at each of the other positions. Since the number of quanta involved in the signal and noise levels received during the resolution periods of duration τ/K is small, these probabilities are more accurately described by Poisson statistics rather than the Gaussian distribution. If $P_A(V_B)$ denotes the cumulative probability that a Poisson distribution with mean A will exceed the value V_B, the probabilities T_N and T_S are given by

$$T_N = P_{N/K}(V_B)$$

$$T_S = P_{(S+N)/K}(V_B)$$

where S is the mean value of the number of quanta in the received signal, and N is the mean number of noise quanta received during the sampling interval. A complete description of the receiver includes specification of the decoding procedure when: (a) no pulse exceeds the threshold during a sampling period and (b) more than one pulse exceeds the threshold during a sampling period. Decoding methods which could be reliably instrumental are to repeat the output of the previous sampling period in the event that no pulse has exceeded the threshold and to decode only the first pulse to exceed the threshold in a sampling period. It may be noted that while the information rate could, theoretically, be improved slightly by more sophisticated decoding techniques, the method described is more representative of what might be accomplished in a technically feasible communication system. Let:

P_i ≡ relative frequency of transmission in the i^{th} position (a priori probability of sending an i)

P_j' = relative frequency of reception in the j^{th} position (a priori probability of receiving a j)

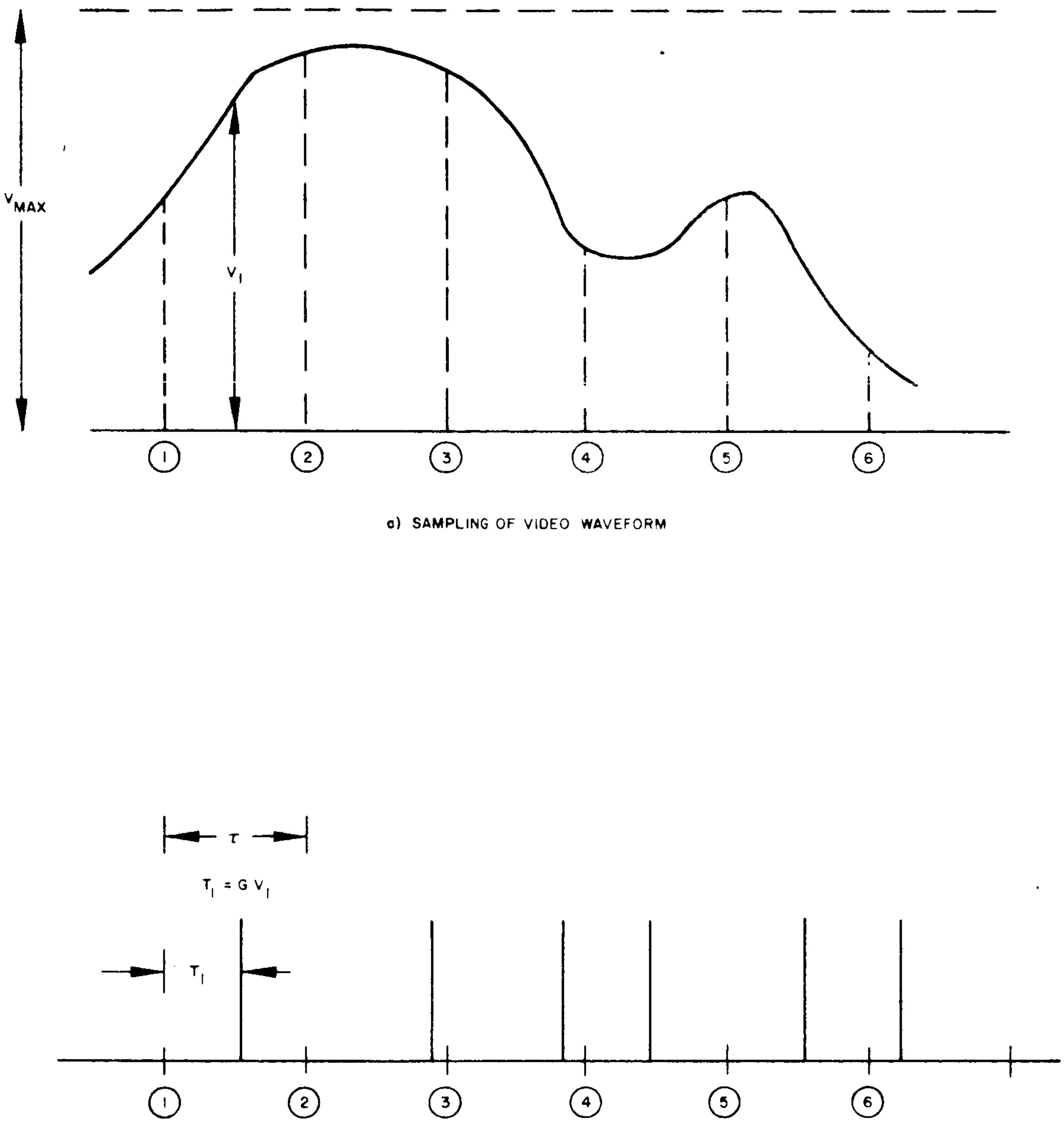
P_{ij} = probability of receiving a j when it is known that i is sent

Then the average number of bits/symbol which can be communicated over the link is:

$$(1) \quad H(S,N) = \sum_{i=1}^{k} P_i \sum_{j=1}^{k} P_{ij} \log_2 P_{ij} - \sum_{j=1}^{k} P_j' \log_2 P_j'$$

This equation[7, 8] embodies the well known information theory statement that the received information is given by the average "entropy" of the signal plus noise (second term) minus the average entropy of the noise alone. Under the

7. S. Goldman, "Information Theory", (Prentice-Hall) 1953, Ch. IV-V.
8. R. C. Jones, "Information Capacity of a Beam of Light", J. Opt. Soc. AM., Vol. 52 No. 5 (May 1962) p. 493.



a) SAMPLING OF VIDEO WAVEFORM

b) TRANSMITTED PULSE POSITION MODULATION SIGNAL

Figure 4. Example of Pulse Position Modulation

decoding procedures defined, the probability P_{ij} is given by:

P_{ij} = (probability that the threshold is exceeded at position j)
x (probability that threshold is not exceeded at a position before j)
+ (probability that the threshold is not exceeded during the interval)
x (probability that the output of the previous interval is a j)

In terms of the signal and noise expectations, T_S and T_N, P_{ij} is given by:

$$P_{ij} = \begin{cases} T_N(1-T_N)^{j-1} + P_j'(1-T_S)(1-T_N)^{K-1} & j < i \\ T_S(1-T_N)^{i-1} + P_i'(1-T_S)(1-T_N)^{K-1} & j = i \\ \dfrac{T_N(1-T_S)(1-T_M)^{j-1}}{(1-T_N)} + P_j'(1-T_S)(1-T_S)^{K-1} & j > i \end{cases}$$

The probabilities are related by

$$(2) \quad P_j' = \sum_{i=1}^{K} P_{ij} P_i = \frac{(1-T_N)^{j-1}}{1-(1-T_S)(1-T_N)^{K-1}} \left\{ T_N + (T_S - T_N) \left[P_j - \frac{T_N}{1-T_N} \sum_{\ell=1}^{j-1} P_\ell \right] \right\}$$

The coding probability P_i is determined by the relative occurrence of light and dark areas in the television picture and the method of compressing the signal amplitude before transmission. In general, when the system is operated at the low post-threshold noise conditions ($T_N \rightarrow 0$) which will be used, the best communication is obtained when P_i is chosen so that the receiver frequency distribution P_j' is uniform, i.e., $P_j' = 1/K$ (j = 1, ... K). For the purposes of computation a transmitter coding will be assumed which produces a variation in P_j' of less than 5 percent over the signal range j = 1 ...K for practical operating parameters and which becomes uniform (optimal) in the case of zero noise.

The coding assumed here,

$$P_i = A\,(1-T_N)^{K-i}$$

is believed typical of the operation which can be achieved in practice. Substitution in Equation 2 yields

$$P_j' = A\,(1-T_N)^{j-1}$$

where $A = \dfrac{T_N}{1-(1-T_N)^K}$ is determined by the normalization of the

probabilities, $\sum_{i=1}^{K} P_i = \sum_{j=1}^{K} P_j' = 1.$

Substituting the expressions for P_i, P_j', and P_{ij} in Equation (1) gives the value of H(S, N) as a function of the probabilities T_S and T_N.

$$(3) \quad H(S,N) = KA\,(1-T_N)^K \left[\frac{\beta - \alpha}{(1-T_N)} + \gamma\right] + \alpha - \gamma \left[1 - T_N\right]^K$$

where

$$\alpha = A/T_N \left[T_N/A + P_o\right] \log_2 \left[T_N/A + P_o\right]$$

$$\beta = A \left[T_S/A + P_o\right] \log_2 \left[T_S/A + P_o\right]$$

$$\gamma = \frac{A}{T_N} \left[\frac{T_N(1-T_S)}{A\,(1-T_N)} + P_o\right] \log_2 \left[\frac{T_N(1-T_S)}{(1-T_N)\,A} + P_o\right]$$

$$P_o = (1-T_S)\,(1-T_N)^{K-1}$$

Two limiting conditions are of particular interest:

LOW NOISE LIMIT: $(T_N \rightarrow 0)$

$$(4) \qquad H(S,N) \rightarrow \left[T_S + \frac{1-T_S}{K}\right] \log_2 \left[KT_S + (1 - T_S)\right] \cong T_S \log_2 KT_S$$

LARGE SIGNAL LIMIT: $(T_S \rightarrow 1)$

$$(5) \qquad H(S,N) \rightarrow - KA\,(1-T_N)^{K-1} \log_2 T_N - \log_2 A$$

Figure 5 shows a plot of H(S, N) for K = 100 for several value of T_N. In general the probability that a noise pulse will exceed the threshold during the sampling period must be kept small; for example, if the probability that the noise exceeds the threshold during one sampling period is 0.3, then the link cannot communicate more than 80 percent of the ideal number of bits/symbol even when arbitrarily large signal levels are used. Similarly the link will suffer if the bias level is set too high relative to the signal since H(S, N) can never be greater than T_S times the ideal value, $\log_2$ (K).

The amount of information to be expected for given signal and noise levels may be obtained by computing T_S and T_N for various bias levels and determining the resultant value of H(S, N) from Figure 5. For given noise level N/K and bias level V_B the value T_N is determined from the Poisson distribution Tables[9]. T_S as a function of the mean (S + N/K) is likewise determined. Figure 6 shows the values of H(S, N) calculated by this method as a function of S for various noise and bias levels. In general, the bias level which gives optimum information efficiency (bits/signal quantum) will vary depending on the value of H at which the system is required to operate. As the desired level of information transmission approaches the maximum capacity of the link, higher bias levels are required. For optimum performance at an information rate in the region of 95 percent of maximum capacity, a bias level should be used for which the noise exceeds the threshold about once in 100 sampling periods.

---

9. "Tables of the Individual and Cumulative Terms of Poisson Distribution", General Electric Co., (Van Nostrand), 1962.

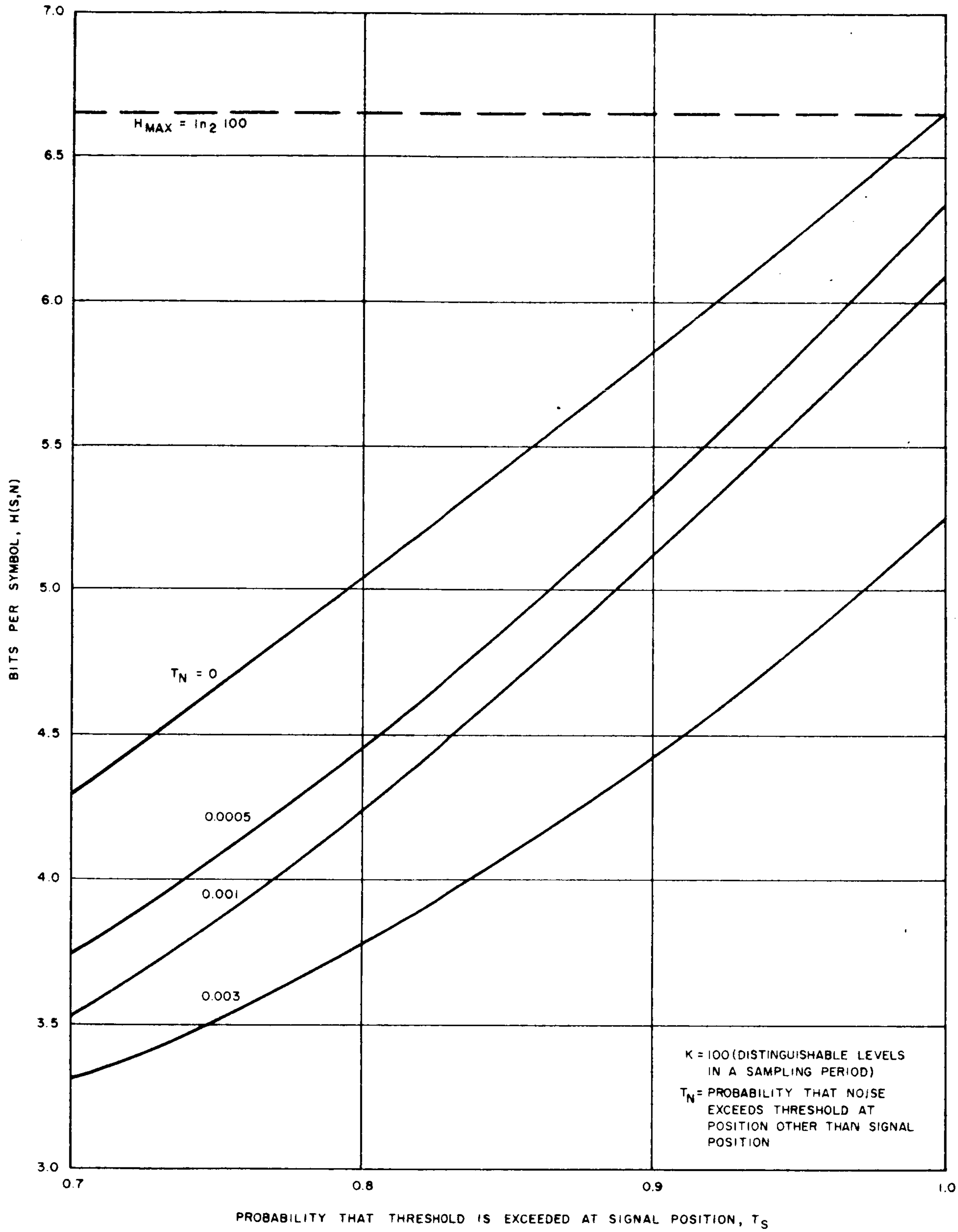


Figure 5. H (S, N) versus Signal Expectation

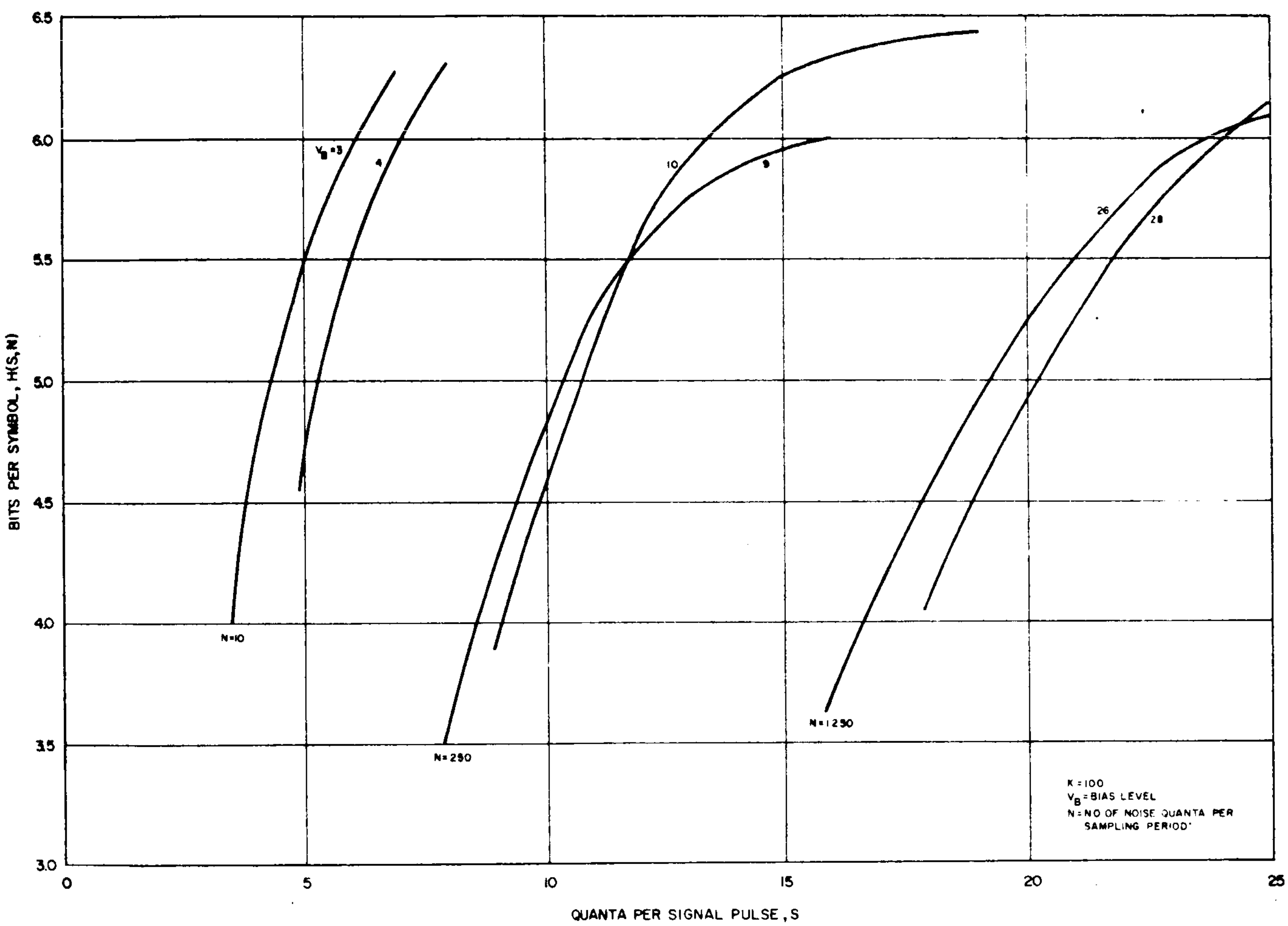


Figure 6. H (S, N) versus Quanta per Signal Pulse

The accurate determination of an optimum bias level can only be done by computing the H vs. S curve for all likely bias levels. Due to the discrete nature of the signal, a bias level must correspond to the detection of an integral number M of quanta in a resolution period of duration τ/K and rejection of all signals of M-1 or less quanta. Thus a bias level of 28.7 is physically equivalent to a bias level of 28.0. This necessity for discontinuous changes in bias level causes some slight irregularities in various of the information rate curves. The problem of selection of optimum thresholds was treated by R. C. Jones[7] in analyzing a similar problem by using a computer to print out results for a large number of bias levels and noting for which conditions optimum information rates are obtained. For the problem at hand, a given bias level will be optimum for an interval of values of H and estimation of the appropriate level can be done fairly accurately without a large number of computations. For a fixed bias level the H vs. S curve rises sharply and forms a knee to approach a limiting value asymptotically. This limit is due to the finite probability of detecting a noise pulse prior to the arrival of the signal and is given by Equation 5. In general, the criterion of 1 noise pulse per 100 sampling periods and $T_S = 0.95$ will place the operating point at the knee of the curve.

Using this rule of thumb, an approximate operating point for a PPM system can be calculated. For a noise current into the receiver giving N noise quanta per sampling period the bias level is determined by the requirement that the probability that a Poisson distribution with mean N/K exceeds V_B should be less than 1/100K. Then a signal level is chosen to give a probability that the signal is detected of approximately 95 percent. Figure 7 shows how the signal required to operate at this point varies with the number of distinguishable positions K in the sampling length for a given noise rate. The signal values defined in this plot locate approximately the knee of the curves of H vs. S (Figure 6). When less signal is available the information rate will decrease rapidly due to insufficient certainty in detecting the signal. Increases above this level will not greatly increase the information rate since the limit $H_o = \log_2 K$ has already been approached.

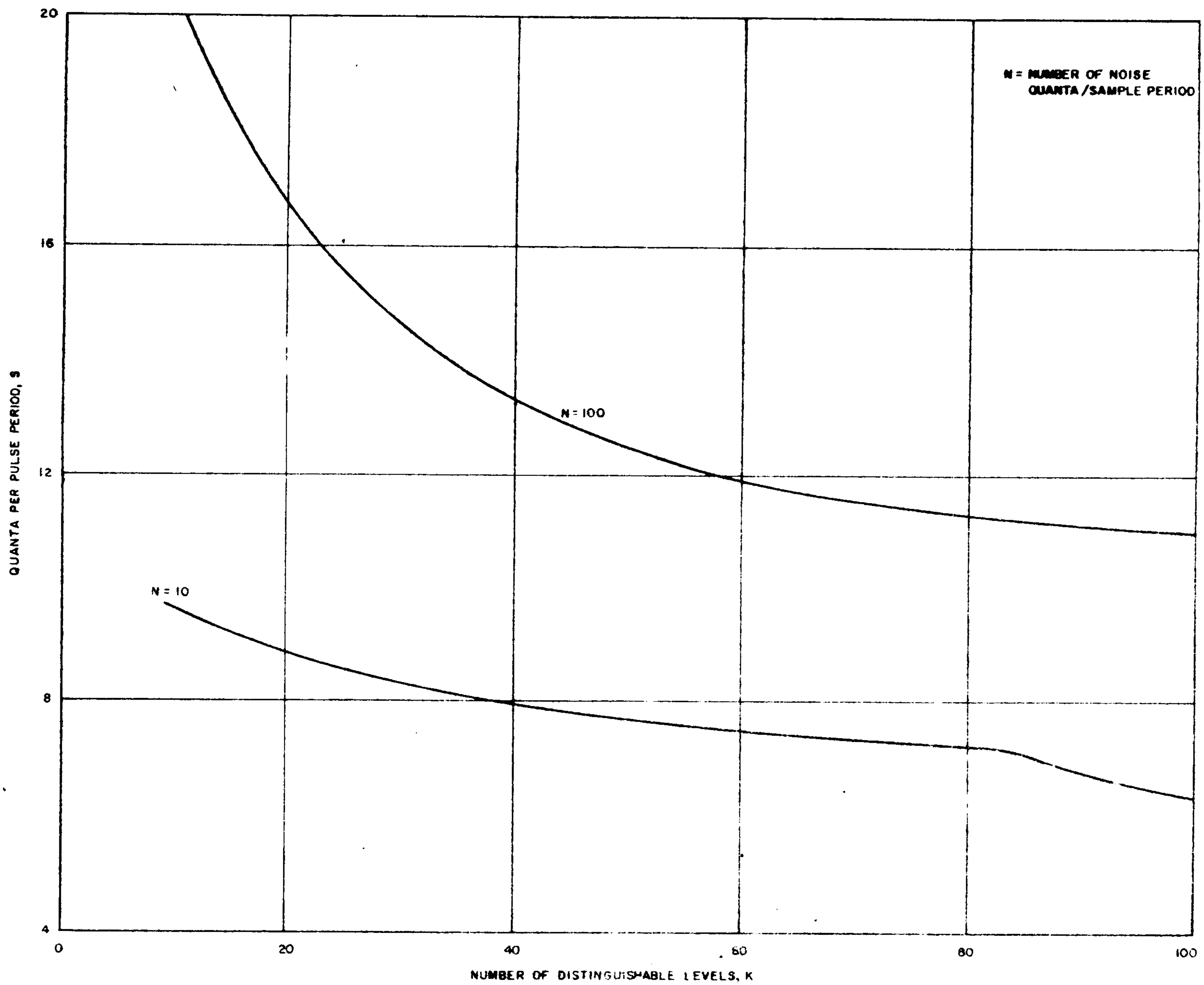


Figure 7. Signal Required to Reach Optimum Operating Point versus Number of Distinguishable Levels

A quantity of interest in evaluating the merits of an optical communication system is the information efficiency, I, which is defined as the average number of bits of information transmitted per signal quantum received. For the PPM system, the theoretical upper limit for the information efficiency increases with K. In a noiseless link a threshold of zero could be used and each transmitted photon would be detected at the correct position in the sampling period. The system could then be operated at a signal level of one quantum per sampling period with an information efficiency $I = H/S = \log_2 K$. This consideration supports the general conclusion that when the noise level is very low and communication bandwidth is not a limiting factor, the PPM technique gives the highest efficiency.

Figure 8 shows the information efficiency for a system with K = 100 and various noise levels when optimum bias levels for operation at the given information rate are used. The information efficiency falls off sharply to zero at $H = H_o$ since the limiting information rate can only be reached by an infinite signal power. One interesting feature of these curves is that the information efficiency remains nearly constant for operation at up to 90 percent (and more) of the maximum information rate. This would indicate that an information rate requirement of H_1 bits/symbol could be met with efficient operation by choosing K only slightly greater than 2^{H_1}.

The amount of information which can be transmitted per signal quantum increases rapidly as the number of distinguishable positions in the sampling period is increased. This is due to the increase in the ideal efficiency ($\log_2 K$) as well as the greater ease with which efficient operation can be obtained (Figure 7) as K increases. The curves of Figure 9 show how the information efficiency varies with K in the presence of noise when the system is operated near the knee of the H vs. S curve. The curves have been normalized to unity at K = 100 to give a result which is only slightly dependent on the external noise level.

When a fixed minimum informate rate is required the signal necessary to provide the rate increases sharply with the rms noise level. The slope of the curve of S vs. $\sqrt{N}$ (Figure 10) increases with increasing H. This is primarily due to the increased bias levels required to minimize S for given H. In general the bias level increases in proportion to $\sqrt{N}$ so that the S vs. $\sqrt{N}$ curve would be expected to be approximately linear.

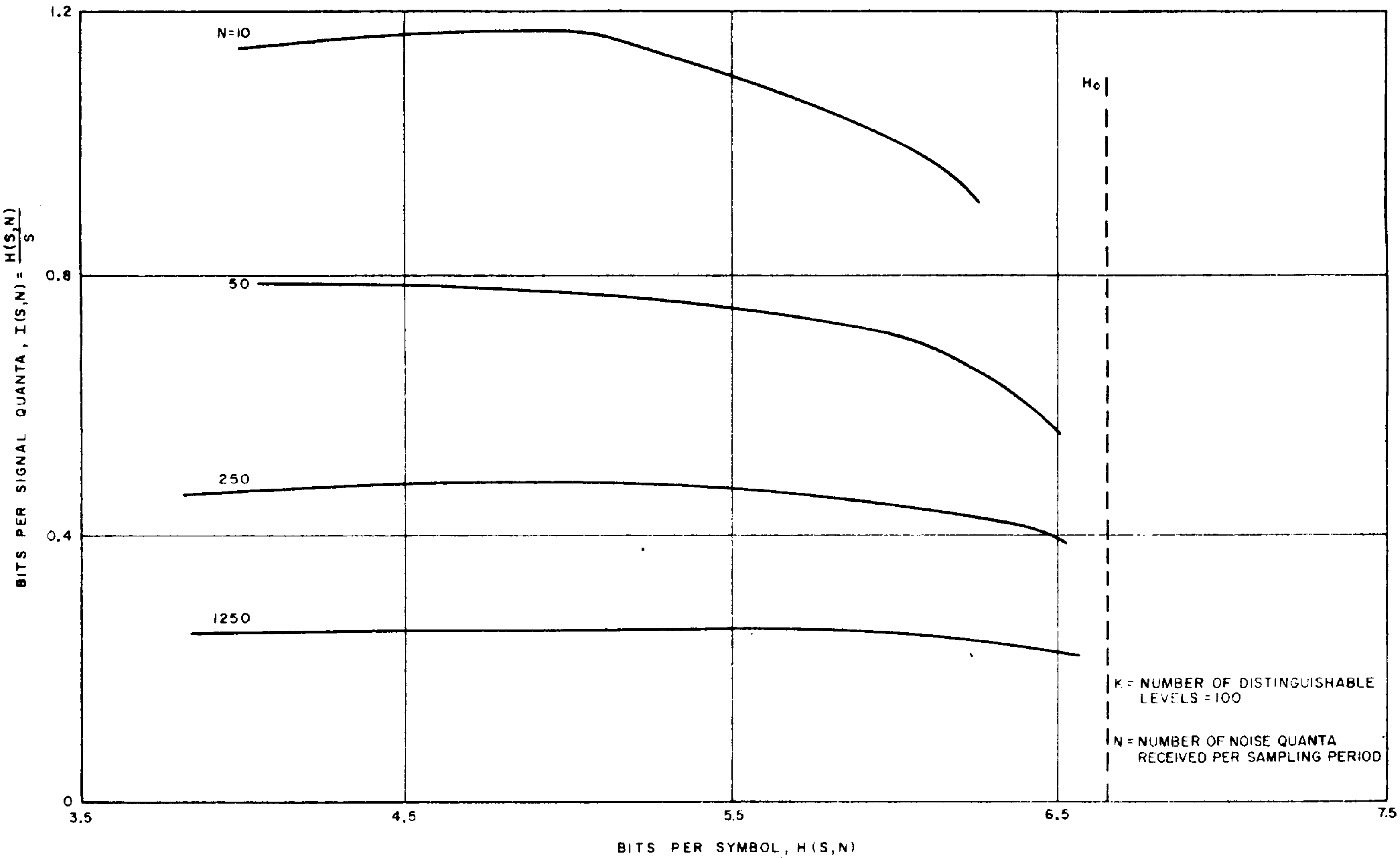


Figure 8. Information Efficiency versus Information per Symbol

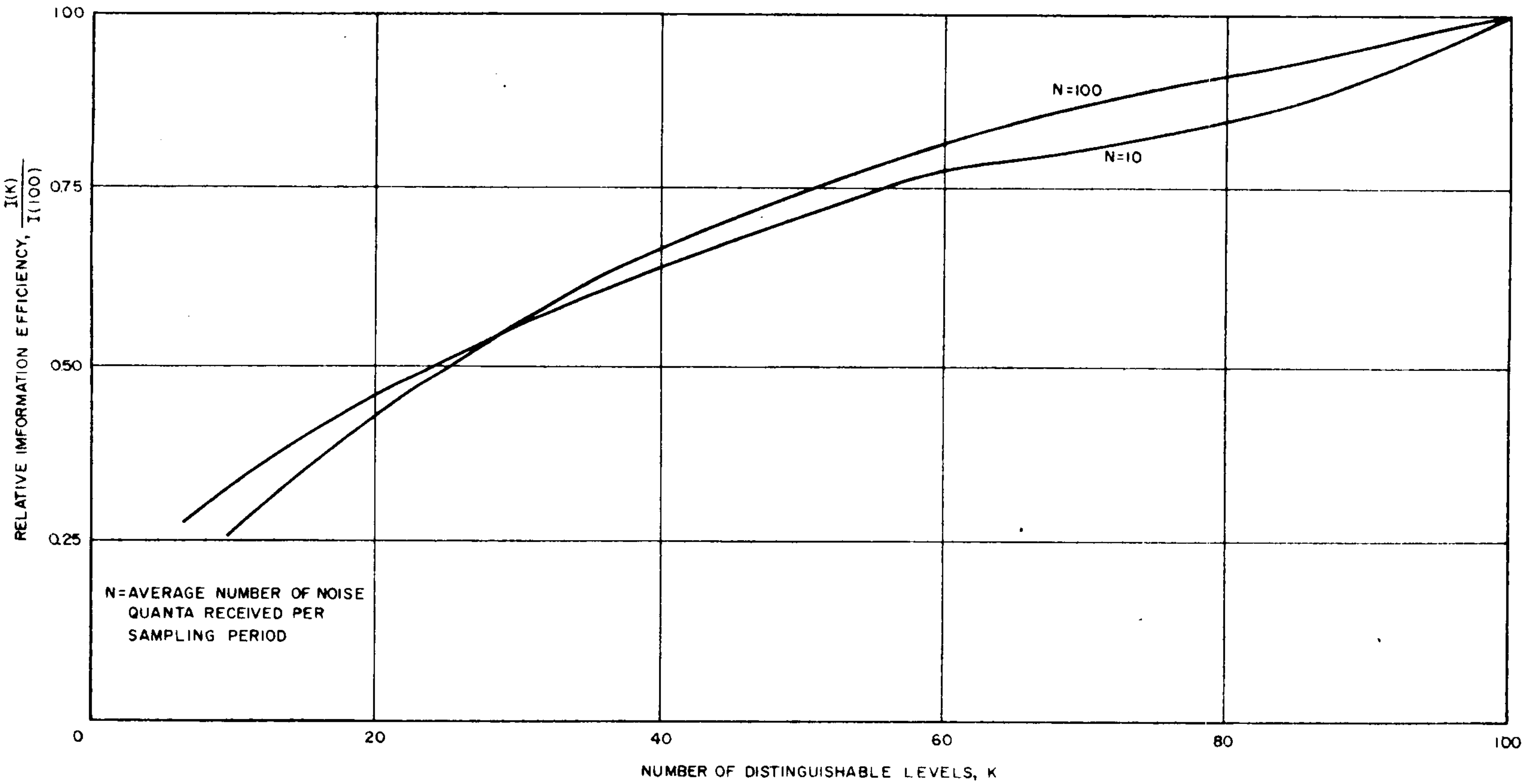


Figure 9. Relative Information Efficiency versus Number of Distinguishable Levels

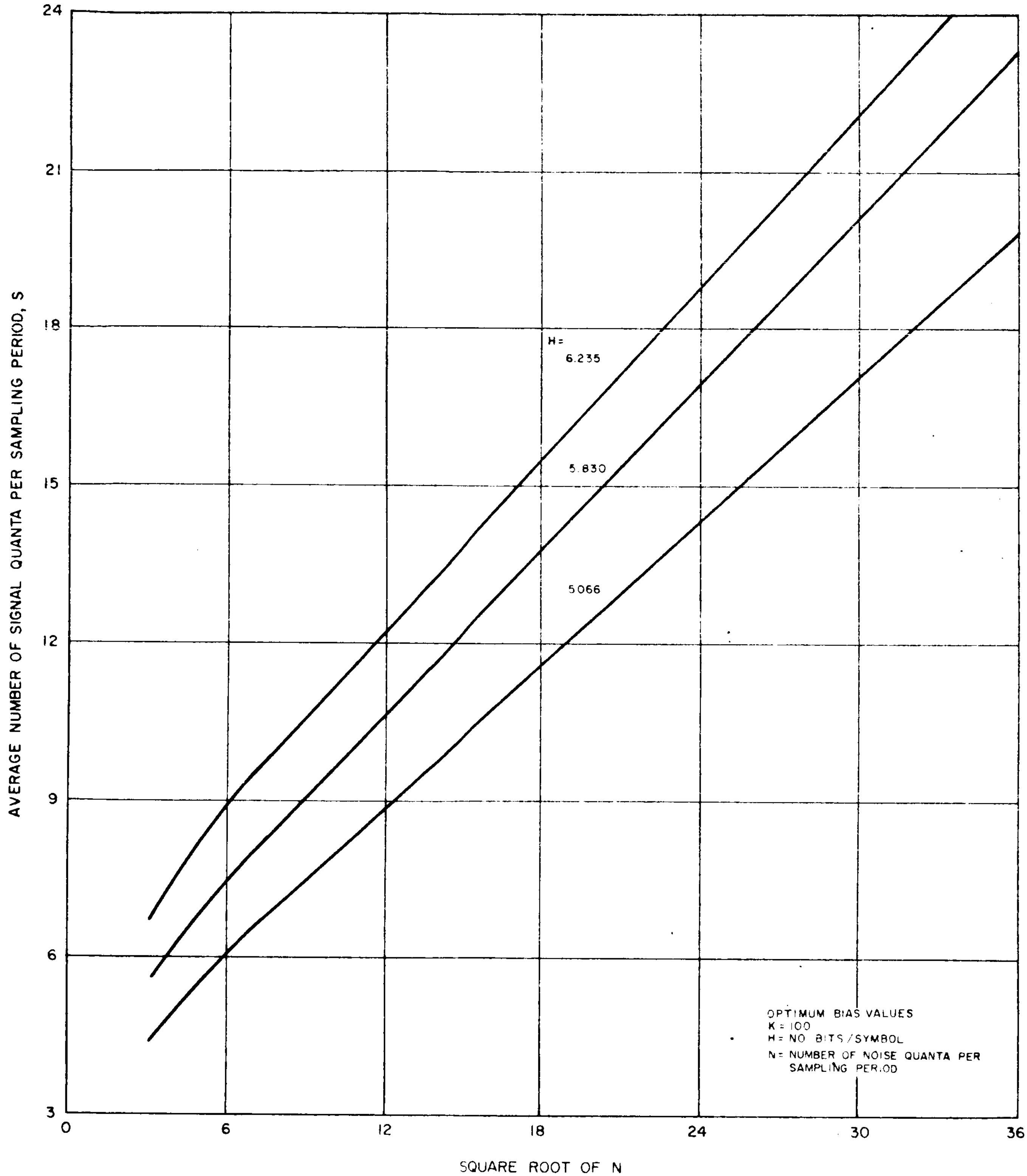


Figure 10. Signal versus RMS Noise

The allowable error rate in the communication link varies with the type of information being transmitted. In the terminology used here the error rate is the average probability that the received symbol is not the transmitted symbol;

$$R_E = \overline{1 - P_{ii}} = \sum_{i=1}^{K} (1-P_{ii}) P_i = 1 - KA \left[T_S - (1-T_S) A (1-T_N)^{K-1} \right] (1-T_N)^{K-1}$$

Figure 11 shows the error rate of the system as a function of S for various noise levels. The eccentricities of these curves are due to discontinous changes introduced when new bias levels are selected.

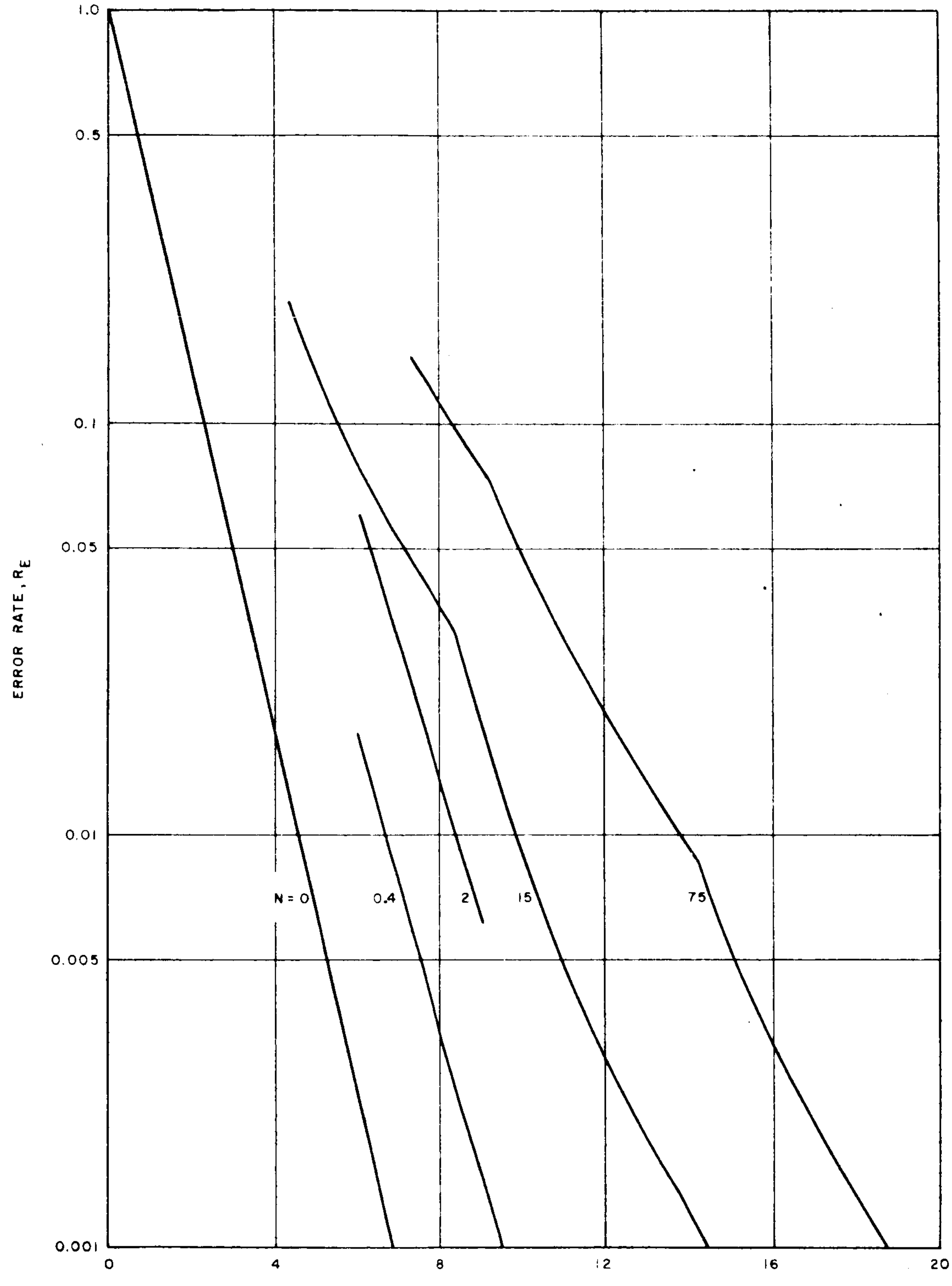


Figure 11. Pulse Position Modulation Error Rate

2.2 INFORMATION RATES OF POLARIZATION MODULATED CONTINUOUS OPTICAL CHANNELS

Elsewhere in this report techniques are discussed for the modulation of an optical beam using a Pockels cell or other modulator. At the receiver the beam is analyzed into its right- and left-hand circular polarization components. The modulation does not alter the total radiant power received, but has the effect of changing the relative amounts of energy in the right- and left-hand polarized channels. When the modulator is adjusted to allow full modulation of the polarization of the laser beam, the received signal intensity in the two channels is:

$$\text{Channel No. 1} \quad \dot{S}_1 = \frac{\dot{S}_o}{2}(1 + X) \text{ quanta/sec}$$

$$\text{Channel No. 2} \quad \dot{S}_2 = \frac{\dot{S}_o}{2}(1 - X) \text{ quanta/sec}$$

where

$\dot{S}_o$ = the quanta current (see previous section) corresponding to reception of the total available signal power at the receiver

X = the modulation which can vary over the range of $-1 \leq X \leq 1$, representing the limits of all to none of the total signal power in one channel.

In the case of a Pockels cell modulator, for example, the quantity X is proportional to the sine of the modulation voltage applied to the cell. By proper signal processing the relation between video waveform and applied modulator voltage can be adjusted so that X could be made, say, directly proportional to the amplitude of the signal to be transmitted.

In the receiver, the outputs of the two channels are differenced to give a signal

$$\dot{S} = \dot{S}_o X$$

with a total range of $2\dot{S}_o$. In the differencing process the individual channel noises are added. The mean square noise level entering the receiver is then given by

$$\dot{S}_o + \dot{N}$$

where

$\dot{N}$ = the average external quantum current entering both channels. Before differencing, the distribution of quanta entering the receiver in each channel is described by a Poisson distribution about the mean value in the channel. Considering sampling periods τ equal $1/2\ B_R$, the mean values received are

$$S_1 + N_1 = \left(\dot{S}_1 + \frac{\dot{N}}{2}\right) \tau$$

and

$$S_2 + N_2 = \left(\dot{S}_2 + \frac{\dot{N}}{2}\right) \tau .$$

The distribution in the differenced channel is given by the cross-correlation function of the two channel distributions. Since this distribution is awkward to handle, for the purposes of analysis, it may be approximated by a Gaussian distribution with mean

$$S = \dot{S}_o X\tau$$

and mean square fluctuation

$$\sigma^2 = S_o + N.$$

Two important types of signal transmission are considered here for an optical channel using a continuously operating laser as signal source. The first is amplitude modulation, in which the amplitude of the instantaneously received difference signal is interpreted as the intended message. The second possibility is pulse code modulation using binary digit coding for the transmitted symbol. For this technique, the sampling period is divided into K segments. In each segment either the full positive signal is present (all the intensity in one polarization channel) or the full negative signal is sent (all the intensity in the other channel). With this type of on-off modulation and using K segments per sampling period, one of 2^K distinct symbols can be sent per sampling period. That is, the channel has a maximum capacity of K bits/symbol. In the AM technique the receiver bandwidth is half the sampling rate while the PCM technique requires a receiver bandwidth of K/2 times the sampling rate for real time communication. The information rates for the two types of transmission, as limited by the available laser beam power, are discussed in the following pages.

A. Amplitude Modulated Continuous Optical Channel

For the AM method, the designation of the transmitted symbol depends on the amplitude of the signal in the difference channel. The noise in the channel is fixed by the differencing process at a rms value of $\sqrt{S_o + N}$. The probability that the amplitude y is received when x is intended is approximately:

$$P_x(y) \simeq \frac{1}{\sqrt{2\pi (S_o + N)}} \exp\left[- \frac{(x-y)^2}{2(S_o + N)}\right], \quad -S_o \leq x \leq S_o$$

where y is an integer. In order to transmit 6-7 bits/symbol, however, the number of signal quanta per pulse must be large so that to a fair approximation the variable y may be regarded as continuous. Then denoting by P(x) and P(y), the a priori probabilities of sending x and receiving y; $H(S_o, N)$ is given by:

$$H(S_o, N) = \int_{-S_o}^{S_o} P(x)\, dx \int_{-S_o}^{S_o} P_x(y) \log_2 P_x(y)\, dy - \int_{-S_o}^{S_o} P(y) \log_2 P(y)\, dy$$

the fluctuation will cause the amplitude received to assume values outside the range $-S_o$ to S_o. However, the large signal range required by the desired information rate will include a very large percentage of the transmitted signals. For the purpose of estimating H the limits of the dy integral of the first term may be replaced by $-\infty$ to $+\infty$. Then:

$$\int_{-S_o}^{S_o} P(x)\,dx \int_{-\infty}^{\infty} \left[-\log_2 \sqrt{2\pi (S_o + N)} - \frac{(x-y)^2}{2(S_o + N)} \log_2 e \right] P_x(y)\,dy = -\frac{1}{2} \log_2 2\pi e\,(S_o + N)$$

The second term is maximum when P(x) is chosen so that P(y) is nearly uniform over the range $-S_o \leq y \leq S_o$. In which case the value

$$-1 \int_{-S_o}^{S_o} \frac{1}{2S_o} \log_2 \frac{1}{2S_o}\,dy = \frac{1}{2} \log_2 4S_o^2$$

could be approached. These considerations lead to the estimate of the information rate for AM transmission in the difference channel;

$$H(S_o, N) \simeq 1/2 \log_2 \frac{4S^2}{2\pi e\,(S_o + N)}$$

this implies that such a system would require several thousand signal quanta to transmit a symbol with 7 bits of information. A considerable increase in efficiency is possible, however, by using the information about the instantaneous fluctuations contained in the sum of the two polarization channels. The average value of the sum of the signals in the two channels is $S_o + N$. When the signal in the difference channel is divided by the signal in the sum channel, the ratio is restricted to within a range varying only slightly with the external noise level. The extreme values of the range, $\pm \frac{S_o}{S_o + N}$ could be adjusted to essentially unity by subtraction of the mean noise level in the sum channel or other correction techniques.

When the transmitted signal quanta in the two polarization channels during the time required to send a symbol is S_1 and S_2 then the instantaneous received values are $R_1 + \delta R_1$ and $R_2 + \delta R_2$ where $R_1 = S_1 + N/2$, $R_2 = S_2 + N/2$ and the probability of a fluctuation $\delta R_1 (\delta R_2)$ is described by a Poisson distribution with mean R_1 (R_2). The value received in the ratio channel is

$$R = \frac{R_1 - R_2 - \delta R_2 + \delta R_1}{R_1 + R_2 + \delta R_2 + \delta R_1} \qquad -1 \leq R \leq 1$$

and the probability of receiving R is the probability that:

$$(R_2 + \delta R_2) - \left(\frac{1-R}{1+R}\right) (R_1 + \delta R_1) = 0$$

which is given by

$$\sum_{j=o}^{\infty} P_{R_2}(j) \; P_{R_1}\left(\frac{1-R}{1+R}\, j\right)$$

where $P_{R_2}(j)$ is the probability of occurrence of the value j in a Poisson distribution of mean R_2. This gives the probability $P_x(R)$ of receiving R when x is sent where $x = \frac{S_2 - S_1}{S_o}$. While the computation of H(S, N) using this precise value of the probability function has not been accomplished, it is believed that the ratio channel results in an order of magnitude improvement in efficiency. However, this improvement still is not sufficient to make the AM method competitive with PPM and PCM techniques. The advantages of AM could only be realized in conditions where bandwidth is strongly limited and ample signal power is available.

B. Information Rate for Polarization Pulse Code Modulation of Continuous Optical Channel

In PCM the video signal is sampled as for the PPM technique discussed in the previous section. The sampling period is divided into K equal segments. The amplitude to be transmitted is then assigned to one of 2^K quantized levels. Each level is then assigned a binary code consisting of a sequency of zeros and ones of length K. A zero is sent in a segment by converting all of the laser beam to, say, right-hand polarized light and a one is sent by converting all the laser beam to left-hand polarized light. The polarization to be transmitted in each segment is then determined by the presence of a zero or a one in that segment of the binary code for the amplitude sampled.

The number of different symbols (sequences) which can be sent is 2^K and the maximum number of bits/symbol transmitted is $\log_2 2^K$ is K. When a segment value (zero or one) is being sent, one of the receiving channels receives a Poisson distribution about the mean $\frac{S_o}{K} + \frac{N}{2K}$, where S_o and N are the signal and noise quanta received per sampling interval. When these two signals are subtracted, the result is a distribution about the value $\frac{S_o}{K}$ and dispersion $\sigma^2 = \frac{S_o}{K} + \frac{N}{K}$. Denoting by $P_A(j)$ the probability of ocurrence of the value j in a Poisson distribution of mean A, the probability of the occurrence of i in the difference channel is:

$$P_d(i) = \begin{cases} \sum_{j=o}^{\infty} P_{\frac{N}{2K}}(j) \; P_{\left(\frac{N}{2K} + \frac{S_o}{K}\right)}(j+i) & i \geq o \\ \sum_{j=o}^{\infty} P_{\frac{N}{2K}}(j-i) \; P_{\frac{N}{2K} + \frac{S_o}{K}}(j) & i \leq o \end{cases}$$

where i must be integral. To a fair approximation, the distribution may be approximated by the Gaussian formula

$$P_d(x) = \frac{1}{\sqrt{2\pi\,(S_o/K + N/K)}} \quad \exp \frac{-(s - S_o)^2}{2(S_o/K + N/K)}$$

The accuracy of the approximation is good for large values of N.

The signal in the difference channel is then taken as a one or zero, depending on the sign of the measured signal. The probability that the determination is correct is given by:

$$T_s = \sum_{i=1}^{\infty} P_d(i) + 1/2\, P_d(o)$$

or in the Gaussian approximation:

$$1 - T_s \simeq \frac{1}{\sqrt{2\pi}} \int_{\frac{S_o}{\sqrt{K(S_o+N)}}}^{\infty} e^{-x^2/2}\, dx$$

In the limit $N \to 0$, T_S may be calculated from

$$T_S = 1 - 1/2\, P_{S_o}(0) = 1 - 1/2\, e^{-S_o}.$$

The expression for the number of bits/symbol which are communicated, Equa. (1), is

$$H(S_o, N) = \sum_{i=1}^{2^K} P_i \sum_{j=1}^{2^K} P_{ij} \log_2 P_{ij} - \sum_{j=1}^{2^K} P_j' \log_2 P_j'$$

Assuming that the a priori probabilities of transmitting different symbols j are the same,

$$P_i = 2^{-K}$$

$$i = 1 \cdot \cdot \cdot \cdot \cdot \cdot \cdot 2^K ,$$

then the a priori probabilities of receiving a j is

$$P_j = 2^{-K}$$

The second term of the expression represents the information/symbol transmitted by the signal and noise

$$- \sum_{j=1}^{2^K} P_j' \log_2 P_j' = K \qquad \text{bits/symbol}$$

The first term in the equation represents the negative contribution of the noise alone. This term may be divided into a series of K terms each giving the contribution of transmission in which the number of errors in segment values (ones or zeros) is the same.

The probability that if i is sent, that the received symbol j contains L errors in segment values is:

$$P_{ij}^{(L)} = (1 - T_S)^L T_S^{(K - L)}, \qquad L = 0 \; . \; . \; . \; . \; K$$

the number of such symbols j is given by $\frac{K!}{(K - L)!\; L!}$

Then

$$\sum_{j=1}^{2^K} P_{ij} \log_2 P_{ij} = \sum_{L=0}^{K} \frac{K!}{(K-L)!\; L!} (1-T_S)^L \; T_S^{(K-L)} \left[L \log_2 (1-T_S) + (K-L) \log_2 T_S \right]$$

$$= K \left[(1-T_S) \log_2 (1-T_S) + T_S \log_2 T_S \right]$$

so that:

$$H(S,N) = K\left[1 + (1-T_S)\log_2(1-T_S) + T_S \log_2 T_S\right] \quad \text{bits/symbol.}$$

Figure 12 gives a plot of $\frac{H(S,N)}{K}$ vs. T_S.

The error rate, defined as the probability that the received symbol is not transmitted signal, is:

$$R_E = (1 - P_{ii}) = 1 - T_S^K$$

Since only error rates on the order of 10^{-2} are of interest, T_S must be very nearly 1 and

$$R_E \simeq K(1 - T_S).$$

In Figure 13 the signal required to transmit with an error rate of 10^{-2} is shown as a function of the square root of the external noise level when 7 bits/symbol (i.e., K = 7) are required.

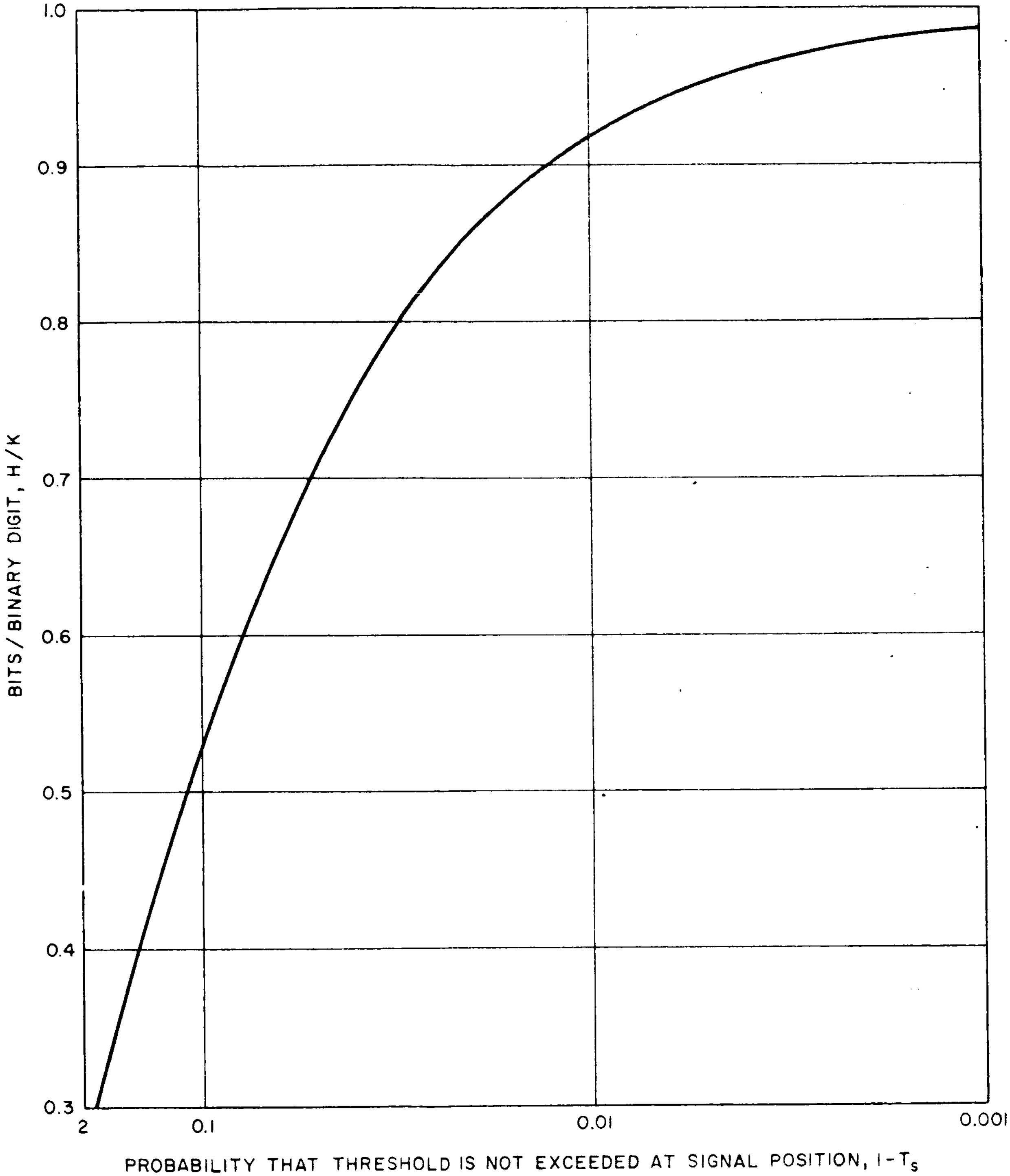


Figure 12. Pulse Code Modulation

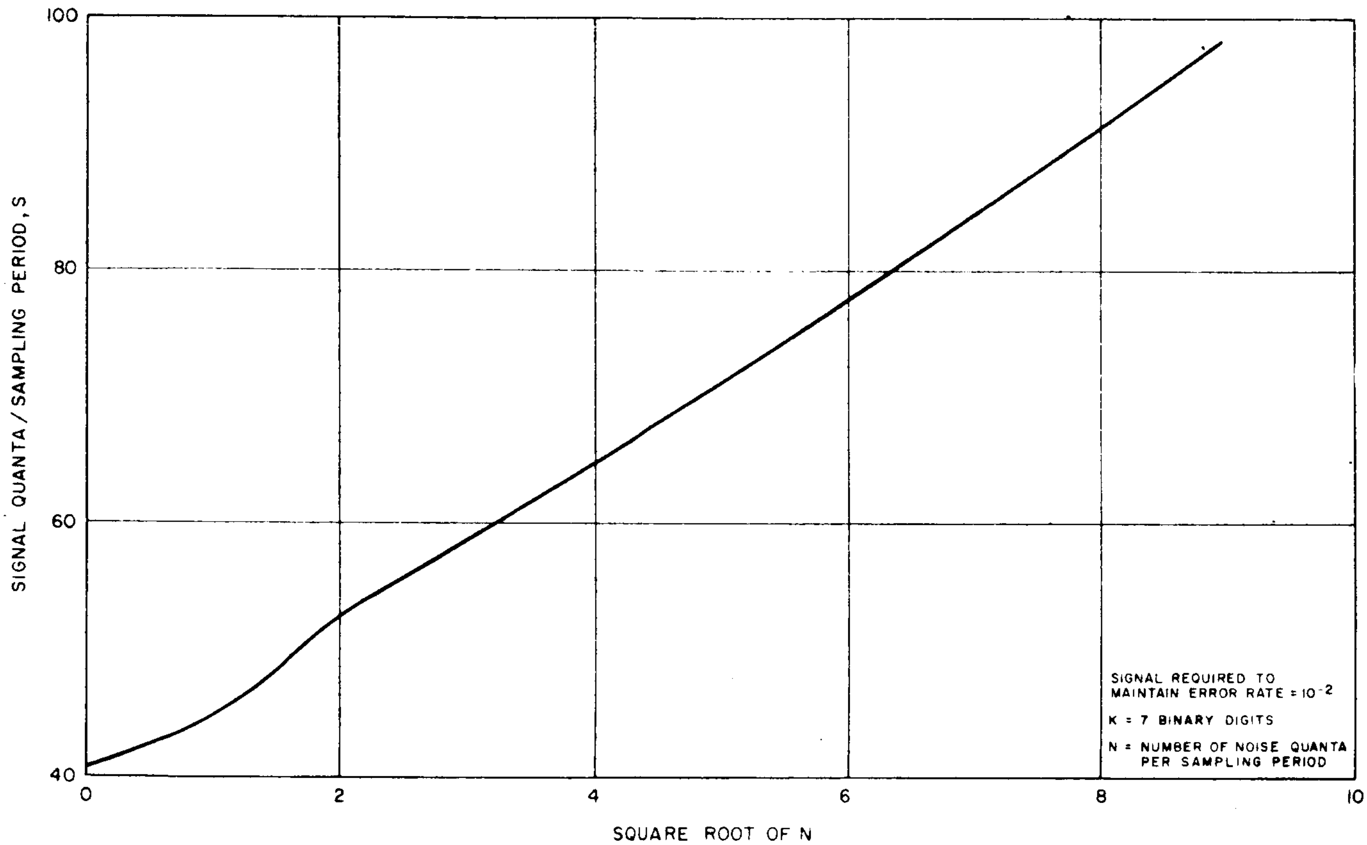


Figure 13. Pulse Code Modulation

3.0 SIGNAL PROCESSING

The type of signal processing to be employed will depend on the types of lasers and detectors available when construction of the system is undertaken. In this section the details of implementation of the schemes discussed from the information theory viewpoint above are presented.

3.1 Pulse Position Modulation

As indicated in Figure 4, the information contained in the amplitude of the video signal at a sample point is converted into the information contained in the time delay of the transmitted pulse from a reference time for that sample period. The proper functional relation between these quantities is suggested by the fact that, for normal levels of illumination, the difference in intensity which the eye can discern is a fixed portion (about two percent) of the total intensity at which the discrimination is to be made. Since the time resolution is independent of the time delay T and since the eye attaches equal significance to proportionate changes in the intensity, which is represented by the video signal, V, the relation

$$T = G \log V$$

is suggested, where G is a gain factor. G should be chosen such that the desired range of intensity fills the available time delay range. If an intensity range of 1000:1 is desired, 100 significant values for T would then present a seven percent significant proportional variation in intensity. The two percent value given earlier applied to large areas viewed for appreciable times, whereas this seven percent value is the instantaneous uncertainty for a single resolution element (sample). Such fluctuations would be essentially unobservable to the eye. Although the value of G and/or video gain would be adjustable to permit optimum transmission of various types of pictures, it will be assumed that they are set to yield a 1000:1 range of intensity and 100 significant levels are needed. By nonlinear preconversion signal compression, the intensity range may be extended although not without sacrifice of fidelity of intensity reproduction at the extremes of the range. Such operations on the video signals do not affect the technique given here unless such optimization reduces the number of significant levels required.

A conceptually simple method for accomplishing the conversion and operating the laser is shown in Figure 14. The operation is as follows: a strong video signal is gated by a clock pulse so that it charges a small capacitor which discharges through a resistor to produce an exponentially decreasing voltage, the initial value of which is the video voltage sample. The time for this voltage to decay to some fixed value is proportional to the logarithm of the video signal. The voltage on the capacitor is used to hold an avalanche transistor in the cut-off condition. Just after the video sample is applied to the small capacitor the B+ voltage is applied via another gate to the energy storage capacitor of the laser and to the avalanche transistor. The decay of the stored video voltage eventually will allow the avalanche transistor to conduct, and the laser will flash. The time delay of the laser pulse is thus proportional to the logarithm of the video voltage.

It is not yet known if suitable voltage levels and impedance values can be obtained without additional isolation and/or amplification, but the present advantage of this circuit, in that only switching operations need be performed in the wide-bandwidth part of the system, should be realizable. Present avalanche transistors switching in one nanosecond and carrying a current of 15 amp. have a switching level stable to 0.002 v., and hence would only need a maximum video voltage a little greater than 0.2 v. if impedances are suitable. Although higher current than can be handled by one of these transistors is actually needed (see the performance section), the probability that the circuits or components eventually developed can still be switched with practical video voltages appears good.

The energy storage capacitor should be of low-inductance, low voltage type. The remainder of the above circuitry, including the laser, can probably be packaged coaxially inside the capacitor. Except for the capacitor and any cooling system needed for the laser, all elements of the system are quite small.

The signal processing in the receiver would be somewhat similar. If an exponentially rising voltage (i.e., a positive exponent) could be generated, e.g., if the exponential build-up of an oscillator could be utilized, direct inverse processes could be employed. However, it appears better to work from the complementary time interval. At the time of arrival of the signal pulse a capacitor voltage decay can be initiated; then the voltage present on the capacitor at the end of the sample period (less any necessary reset period) represents the transmitted signal. A clock pulse gates out the video by sampling this exponentially decaying capacitor voltage at the correct instant. This technique is illustrated in Figure 15.

3.2 Polarization Amplitude Modulation

In polarization amplitude modulation the state of polarization of the transmitted beam is proportional to the video signal. Polarized light can be separated into not more than two completely distinct portions. Typically, these two are orthogonal linear polarizations or right- and left-hand circular polarization; although orthogonal elliptic polarizations are also possible, they are representable as superpositions of the above classes. It is not possible to polarize a beam so that it can be divided into, say, four distinct portions such that any portion will be present in any desired amount with only the restriction that the sum of the parts equal the whole; if this is attempted it will always be found that relations between the classes are such that no additional information can be conveyed by the two additional portions. Therefore, in polarization modulation the transmitted signal can only be conveyed by the relative amounts of the two distinct polarizations.

The state of polarization of the beam at reception will be determined by separating the two significant polarizations by means of a polarizing prism. Each polarization may then be detected separately. Light from the daytime sky is partially unpolarized and partially linearly polarized. It therefore might be detected unequally by the two channels if linear polarization were transmitted. Therefore, it is desirable to transmit circularly polarized light. This can be converted to linearly polarized light for separation by the polarizing prism by use of a quarter-wave plate in the receiver. Left-hand circular polarization will produce a linear polarization orthogonal to that produced by right-hand circular polarization; the particular polarization directions produced are determined by the orientation of the axis of the quarter-wave plate. An inverse process can be used to produce circularly polarized light at transmission if it is more convenient to produce linearly polarized light in the modulator.

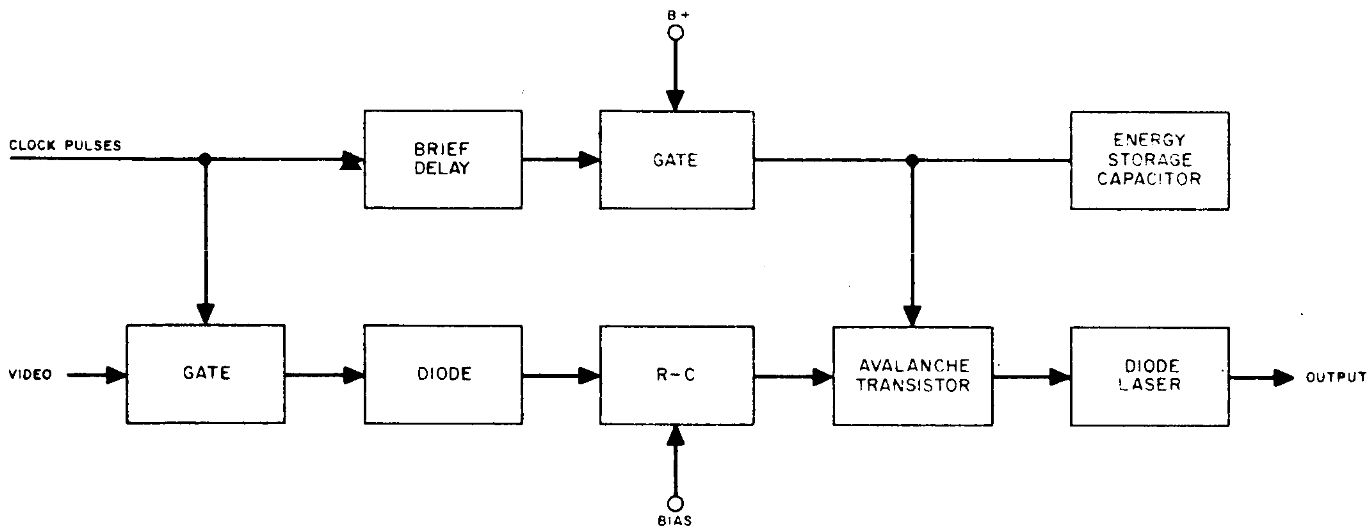


Figure 14. Method of Generating a Pulse Position Modulated Lasser Beam

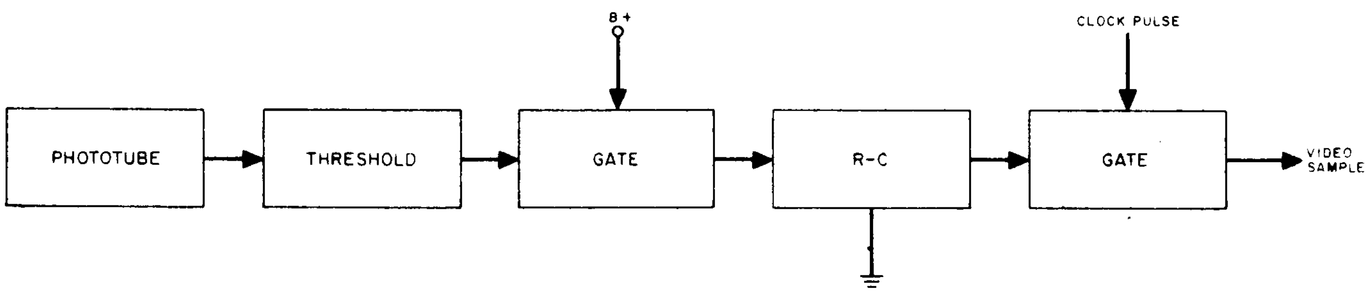


Figure 15. Method of Recovering Video from the Pulse Position Modulated Beam

There are no special techniques essential for modulation and demodulation of this type of signal. A powerful video signal will probably be needed to modulate the polarization by means of electro-optical devices. The receiver need only obtain the difference between the signals from the detectors receiving the two different polarizations. Nonlinear operations may be used to normalize this difference, allowing for statistical fluctuations in the sum of the two signals, amplitude scintillation, and sky light. Such operations have not been worked out in detail because the optimum correction is difficult to determine. The simple difference between channels lies within a range of twice the sum of the channels (minus to plus); it may therefore be normalized to a fixed range, which is a characteristic of the transmitted values, by division by the sum. However, the sum includes some sky light, and hence this process would produce a reduced contrast in the picture. If adjustment of contrast is made, some values outside the dynamic range of the channel are produced. This may provide the most acceptable picture, although in principle a more correct intensity reproduction could be made by a nonlinear compression of the extreme values.

3.3 Polarization Digital Modulation

Polarization digital modulation is much the same as polarization amplitude modulation. The sample interval is broken up into perhaps seven pulse intervals which are used to transmit a binary number characterizing the intensity at the sample point. Ones and zeros correspond to transmission of one polarization or the other. Reception is simple in that it is only necessary to determine the sign of the difference signal to determine whether a one or a zero has been received. If a zero difference in photon count is obtained, system internal noise will cause either positive or negative output with equal probability. Clock synchronism may be obtained by either transmitting synchronization pulses at the end of each line of the picture or by use of some sort of a phase-lock circuit nulling the randomly generated second harmonic ($1/\tau$ instead of $1/(2\tau)$) in the receiver. This transmission technique permits use of much lower signal-to-noise ratios than polarization amplitude modulation, although it does require rapid analog-to-digital conversion and more precise clock synchronization. Slightly more than 100 distinguishable levels are needed unless fluctuation in the levels is introduced so that the eye does not have time to distinguish the quantization steps. For example, levels could be shifted one-half step every other frame, and persistence of vision would then provide approximately the equivalent of twice as many levels. Synchronization in this coding shift is necessary to keep true relations between transmitted and received values.

3.4 Low-Data-Rate Transmission

Since the equipment necessary for this system to transmit real-time standard television data rate 5×10^7 n mi may not be developed soon, it is important to consider means for reducing this data rate. One means is simply to record images and then transmit over a longer period of time. However, real-time transmission may be desired. In that case, it is possible to reduce the data rate by taking advantage of the spatial and/or temporal redundancy of the picture. In ordinary television transmission it is assumed that every resolution element is independent of its neighbors and of its previous values; i.e., it is assumed that the intensity for every resolution element must be transmitted each time it is to be displayed. This assumption simplifies the receiver circuitry, since then no storage or correlation is needed. However, television always contains some such redundancy. Temporal redundancy exists in that only a few objects in the picture are moving, and hence it is wasteful to transmit the same value repeatedly for every element. In fact, even the scan rate is determined more from the point of flicker reduction, obtainable by simply multiple display of the same image, than from the requirement of apparent continuity of motion. Spatial redundancy exists in the presence of broad areas over which the intensity is slowly varying; these areas are usually separated by sharp boundaries, which represent the high-frequency portion of the video. Separate transmission of edge data and low-frequency data permits substantial reduction in the data rate.

Work is proceeding at Hughes on two programs of television bandwidth reduction, one concerned primarily with temporal redundancy reduction and the other spatial.

unnecessary at the moment to relate the signal processing approach here to the details of these systems, it can be pointed out that the analog approach to removal of temporal redundancy can be readily represented as a simple reduction in video bandwidth. Points are merely sampled less frequently, and such samples can be used to reconstitute a video of reduced bandwidth, which can then be processed in the receiver at the earth station to provide the desired full video bandwidth available by use of the previous video samples. Therefore signal processing in transmission can be done just as though the transmission were slowed down by a factor of five, and the foregoing theory and design approaches can be used directly. The digital output resulting from the present Hughes approach to reduction of spatial redundancy leads directly to digital transmission, although it should be possible to convert it to an analog signal output to permit use of the PPM transmission technique. Such details will not be considered here.

4.0 DETAILS OF THE DESIGN EXAMPLE

The design example is intended to represent one specific embodiment of a set of the design concepts discussed in the foregoing sections. As improved knowledge is obtained of both the nature of the mission and the state of the art at the time of final selection of a design for that mission, one or another design approach may appear preferable. It is hoped that the following will be indicative of the general features of the system finally adopted.

4.1 Mounting to the Spacecraft

Continuous rotation of the spacecraft may be desired either to provide synthetic gravity or to permit one side of the spacecraft to continue to face a planet about which it is orbiting. It is therefore desirable to provide a continuous rotation of the gimbal for the telescope of the optical communication system. This suggests a gimbal of the alt-azimuth type, i.e., that a continuous-rotation "azimuth" axis be established perpendicular to the side of the spacecraft on which the system is mounted. If this axis is also parallel to the spacecraft axis of rotation, a fixed object can be kept continuously in view in spite of the spacecraft rotation. The mounting shown in Figure 16 has elevation angle limits of -32° to $+80^{\circ}$; it therefore can cover approximately 75 percent of the full sphere.

For protection during launch and at other times when the system is not in use, it would be stored inside a closed, airtight chamber. This chamber has a door opening into the spacecraft to permit the crew to do any necessary maintenance, e.g., lubrication, cleaning or adjustment of optics, or replacement of faulty components. The gimbal structure swings out of the enclosure, opening the lid as it goes; the lid latch is a circumferential ring with a multi-pitch thread similar to that used on quarter-turn jar lids. The screw drive to swing the gimbal up and out is similar to that used on aircraft landing flaps. When the gimbal is fully out or fully in the drive stops but remains under stress so that the structure is held firmly against a stop.

The drive for the main gimbals uses direct-acting torque motors. A double-integration servo is used, probably with rate gyro feedback for stability. The error signal is developed by sensing the deflection of the isolation gimbal. The main gimbal therefore serves only to minimize

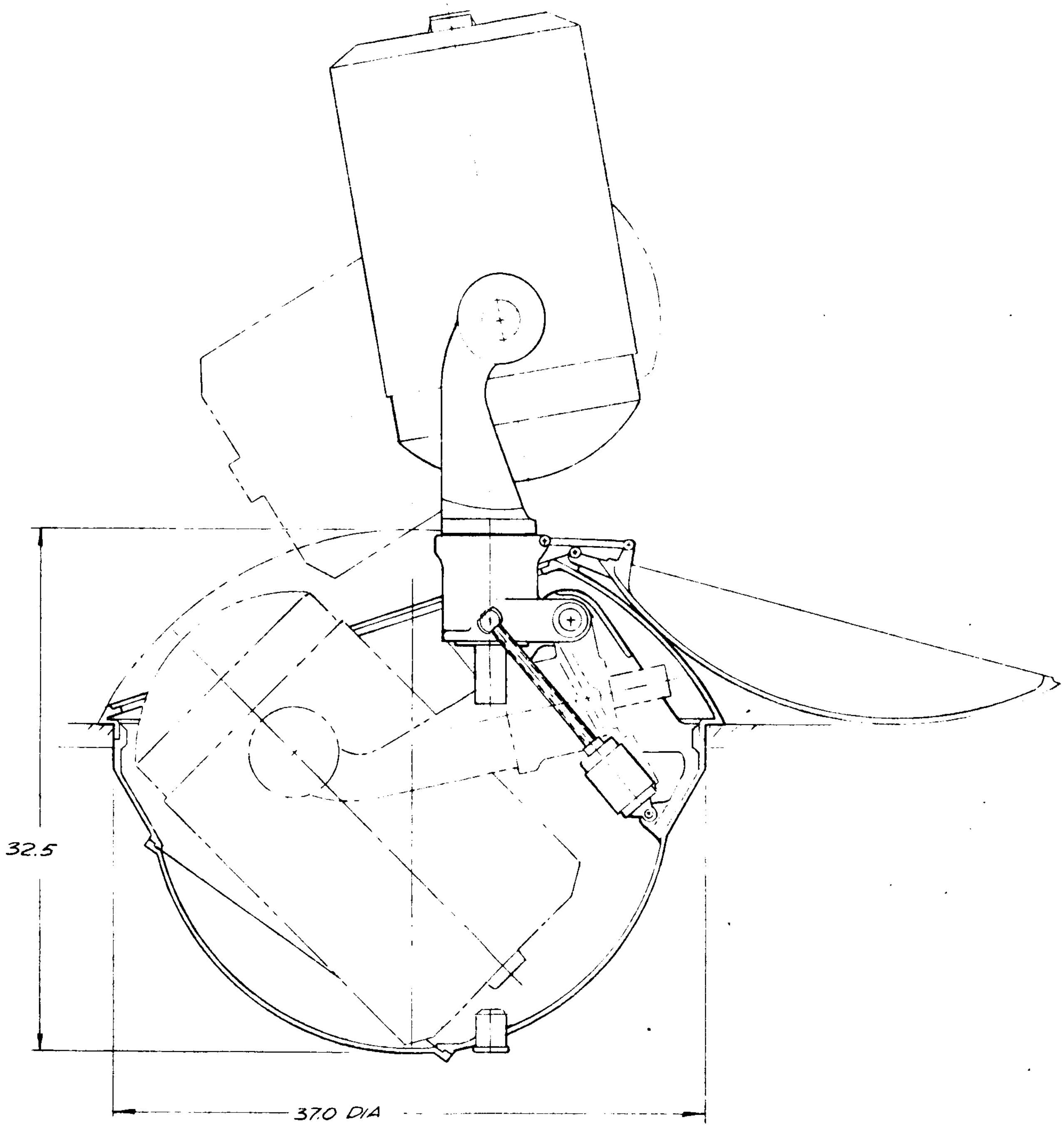


Figure 16. Spacecraft Mounting

deflection of the isolation gimbal.

As indicated in the illustrations, practically all the mechanism is on the main gimbal. The only active element on the isolation gimbal is a focusing device at the secondary mirror. Electrical power, control, and signal connections must go through slip rings or other rotary connections at the azimuth pivot, but can simply flex for the elevation pivot and the pivot for swinging the mechanism in and out of the enclosure.

No special provisions for handling sunlight coming into the telescope tube are indicated; these mechanisms, if needed, would be specifically designed for the conditions to be encountered during a particular mission.

When the telescope is in the enclosure it may be completely dismounted via the internal access door. It is also possible to simply point it into the interior of the spacecraft through this opening and thereby perform optical tests on it with all portions of it operating.

4.2 Primary Optical System

The primary optical system, shown in Figure 17, consists of a two-mirror, f/50 Cassegrainian configuration employed as a telephoto lens. Since the objective diameter is 12 inches, the focal length is 50 feet or 600 inches. For proper pointing of the 0.005 mrad-diameter transmitter beam the rms pointing error must be reduced to less than 0.001 mrad or 0.0006 inch at the focal surface. The main purpose of having an f/50 system rather than, say, an f/10 system is to give a sufficiently large image that this tolerance is reasonable. Once the transmitter and receiver beams are separated absolute values must be maintained for boresight, and hence larger dimensional tolerances are desirable. The most difficult such tolerance remaining is that for focus of the primary optical system. A simple means by which that can be checked during operation is provided.

4.3 Isolation Gimbal

The isolation gimbal pivots about the focus. Therefore pivoting of the main gimbal does not affect pointing, although translation of it with respect to the isolation gimbal would. Angular motion of the main gimbal does affect the portion of the aperture which the laser illuminates; a deflection of one

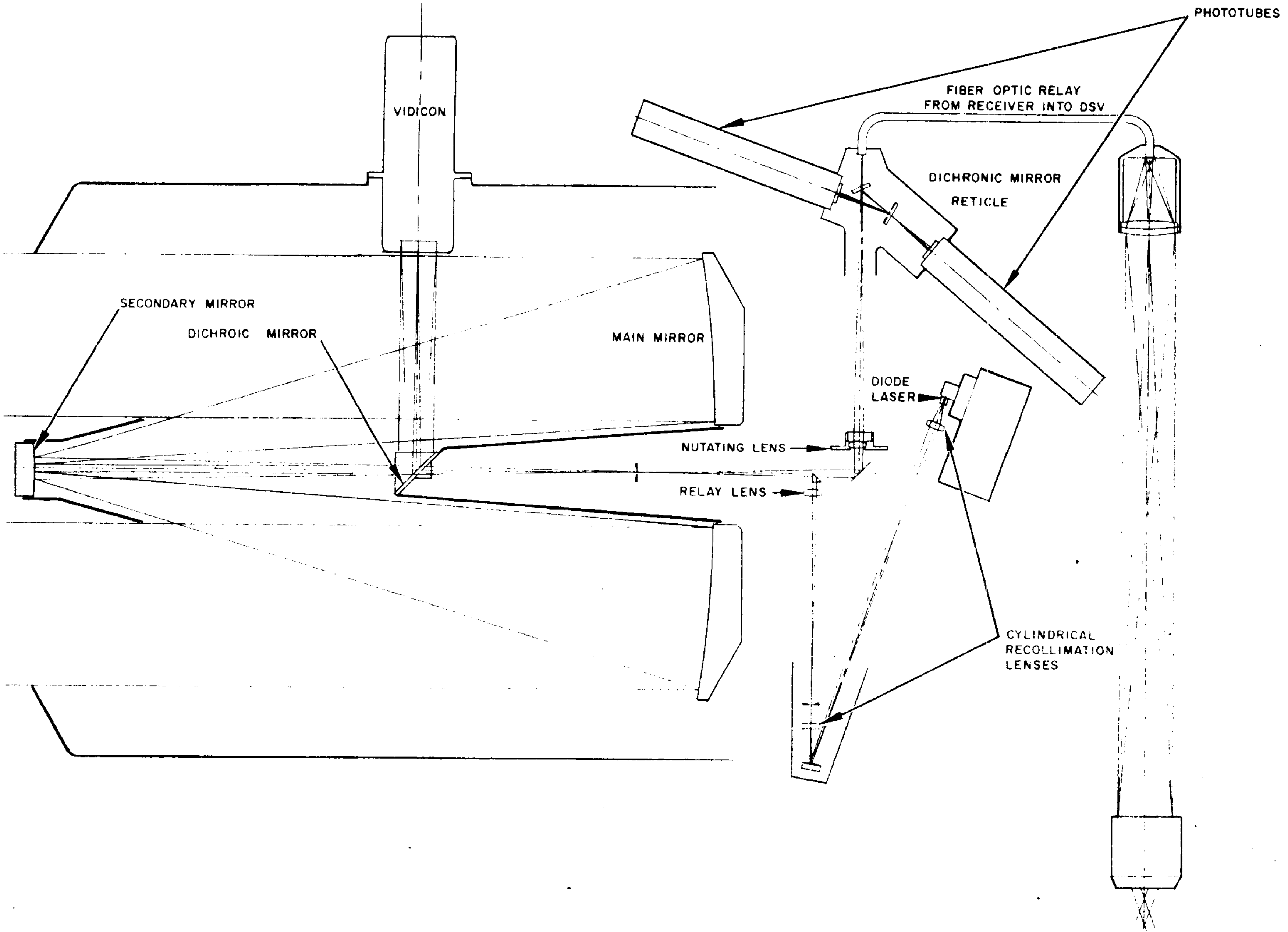


Figure 17. Primary Optical System

mrad corresponds to a shift of 0.6 inch across the main mirror. Angular deflections of the isolation gimbal pivot must therefore be held to less than one mrad; a 0.001 rms error limit of 0.2 mrad should be obtainable.

Not only should transverse linear deflections of the isolation gimbal be held to 0.0006 inch rms to obtain mrad rms error for pointing the 0.005-mrad-diameter transmitter beam, but longitudinal deflection should be less than 0.01 inch to prevent detectable shift of focus, although system operation would not be seriously degraded by a shift of 0.1 inch.

The correct type of stiffness to meet these requirements is provided by a flexure pivot mechanism consisting of an annular plate of thin metal, bonded at the inner diameter to the main gimbal and at the outer diameter to the structure carrying the primary optical system. (See Figure 18). The ratio of longitudinal to angular stiffness is maximized by increasing the proportionate width of the annulus, but there is little advantage in going beyond a ratio of inner to outer diameter equal to 0.5. The angular stiffness required is approximately 10 gm-cm per degree, and hence the pivot will be fairly delicate, unable to support the weight of the isolated mechanism during high accelerations. However, it will be sufficiently compliant longitudinally that it can simply be made to bottom on a solid support which will then transmit the forces.

Torquing of the isolation gimbal will be by means of small solenoids acting on tiny permanent magnets. The force will thus be made independent of position and velocity of the main gimbal. Only small forces will ordinarily be needed. Larger forces for gross motions (e.g., when the telescope is being extended and put into operation or when its function is being changed) can be accomplished simply by letting the main gimbal go to the end of the free motion of the isolation gimbal and hence carry it along. Error signals for tracking will be generated by sensing the position of the earth beacon. Rate feedback for stability can probably be taken from the main gimbal because the servo for that drive will be much faster, and hence the delay in rate feedback will be unimportant. Perturbations introduced into the main gimbal servo loop will be coupled into the isolation gimbal drive, but further study would be required to determine if they would be serious; it does appear that at least such a system could operate stably. Perhaps the spring coupling and the rate gyro coupling could offset each other.

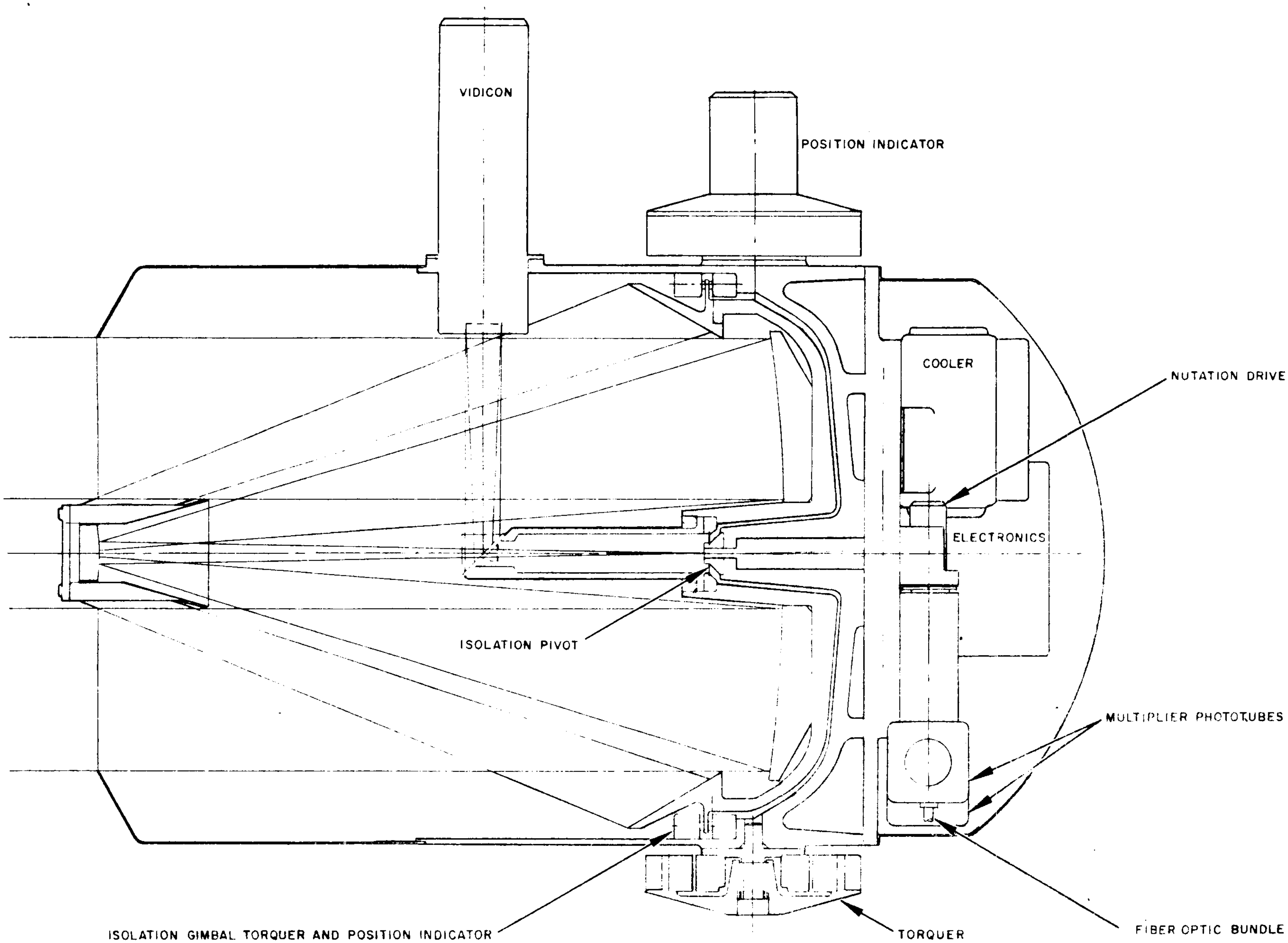


Figure 18. Flexure Pivot Mechanism

4.4 Divisional Optical Systems

The first division of the received light energy removes all the visible light except red light of wavelength greater than about 0.6µ and reflects it to a vidicon which thus provides a television camera with a field of view of about 1.6 mrad and a resolution of about 0.003 mrad. This television camera performs two important functions. Firstly, it can be used to generate error signals for tracking the earth when the system is not locked onto a beacon, and secondly, it can be used to provide images of the earth and other planets when the spacecraft is in transit, and to examine the surface of a planet in detail when the spacecraft is orbiting it. These second functions provide an argument for providing two optical communication systems; not only is redundancy provided, but these telescopic views of other planets can be relayed to earth.

The next division of the received light energy occurs when part of the beam, representing a portion of the aperture, is extracted by a mirror and sent into the transmitter. It is simply wasted there, of course; the reason for the mirror is to permit insertion of the transmitter beam, and it simply forms an aperture blockage for the receiver. At this point the transmitter beam is converging to the same focus the beam for the receiver is diverging from; throughout this area behind focus stability of optical elements must be adequate to maintain this condition without overly frequent adjustments.

The beam to the receiver passes through the combined relay lens and nutation mechanism and goes to another dichroic mirror which separates the received signal from most of the remaining red light. The latter is focused on the end of a flexible fiber optic bundle which relays the image across the elevation pivot and down the axis of the azimuth pivot, where a lens projects it (still rotating with the azimuth gimbal) across to a receiver on the inner wall of the system enclosure. It is thereby possible to verify the relation between what is viewed on the television system and what the receiver is viewing. The fiber optic bundle is located very close to the receiver focus and only solid, non-adjustable parts are depended upon to maintain correlation between what it views and what the receiver views. It would be possible to let some of the light from the beacon go to this fiber optic bundle so that the beacon could readily be seen visually; however, enough beacon light will pass that the beacon can be seen when it is on the night hemisphere of the earth, and that should be adequate. During the day identification of geographic features will suffice. It will be possible to view with the nutation drive stopped, and that would be best for checking focus, but it will also be possible to remove the effect of nutation (for viewing) simply by introducing a compensating nutation. This will also blur out the mosaic of fibers in the bundle, although with the type of viewing employed here they may be practically invisible anyway.

The image for the receiver falls on a transmission-reflection-absorption reticle pattern. The simplest such patterns have an aluminized spot surrounded by a clear area which, in turn, is surrounded by a black, opaque mask. The light reflected from the spot falls on one detector and that transmitted by the clear area falls on another. The beacon signals have a high-frequency carrier to distinguish them from other light sources on earth. This carrier is first detected, and then the difference in the detector signals at the nutation frequency represents the tracking error; if there is no error, the image moves around the edge of the aluminized disk. The phase of the signal tells the direction of the error. The error signal may be passed through a resolver synchronized with the nutation drive, or it may be mixed with electronic signals related to a nonrotational nutation drive as sine and cosine of the phase angle of the displacement with respect to one of the reference axes. The sum of the two detector signals carries the voice and/or other command information. Note that unless tracking disturbances are serious and/or beacon signals are weak, one detector suffices for both outputs; i. e., when errors are small and noise is not a problem each detector would give half the output signal when compared with zero.

Each phototube has a narrow-band interference filter placed over it. These filters must be approximately perpendicular to the beam they are to filter. The divergence of the f/100 beam present at this point in the system is not toublesome, but the angle at the dichroic mirror sending light to the fiber optic bundle is too great to allow filtering an f/100 beam there. Except for the doubling of the number of narrow-band filters actually needed, the arrangement is not wasteful because secondary filters are always needed with narrow-band filters to remove secondary transmission maxima. Narrow-band filters must also be controlled in temperature. A Peltier heater-cooler would probably be the best approach to this requirement. Another possibility is to recollimate the beam so that a slight angular deviation is permissible and then to tilt the filter, thereby producing a wavelength shift to offset that caused by temperature change.

4.5 The Laser

If the transmission of a supposedly random (spatially and temporally) picture at regular television rates for a distance of 5×10^7 nautical miles is required, we find that no present-day laser is adequate. Even when the data rate is reduced by a factor of five by taking advantage of the redundancy normally present in a television picture, there is still no "off the shelf" laser for this task. For the design example it will be most fruitful to consider the electro-luminescent diode laser operated in a pulse position modulation system. If we assume laser performance at $0.63\,\mu$ approximately equivalent to that obtained at $0.85\,\mu$, and that five nanosecond pulses at 2×10^6 per second are available and can be detected efficiently, a desirably low-power system can be hypothesized. Since none of the suppositions appear to be very far from present developments, and since alternative lasers may well be developed, this seems a reasonable way to proceed toward the desired system. The pulse position modulation system is theoretically a very efficient type of modulation, and the electroluminescent diodes are theoretically capable of high efficiency.

Because of its normally elongated shape, the laser is initially collimated by crossed cylindrical lenses, one placed close to the laser to form a magnified image and the other placed close to the image to form a minified image at the same focus so that the apparent cross section of the emitter is stretched in one direction and compressed in the other. A normal relay lens system follows: It is divided into two parts so that focus and boresight (including lead-angle) adjustment may readily be made. The feasibility of this collimation depends on how closely the diode lasers approach the diffraction limit. For present devices this would correspond to a 0.1 degree by 20 degree beam, approximately.

The chief disadvantage of the diode lasers is that they probably will have to be cooled. However, in spite of this, their over-all efficiency (including cooling power) will still be good. A 100-watt Stirling-cycle cooler will deliver 10 watts of cooling at 80°K. It would weigh only 3 to 5 pounds, including a pump to circulate a cooling fluid about the main gimbal for removal of its generated heat. Radiation from the main gimbal itself would be adequate to dissipate this and all other power dissipations. The efficiency of the laser is quite important to the cooling system design because power dissipated in losses in the laser must be removed as heat by the cooling system.

Another important aspect of the diode lasers is that each electron passing through them produces at most one photon in the output. From the values for the charge of the electron and the energy of a photon of 0.63- light it may then be shown that the diode emits at most two watts per ampere. Therefore, if high peak powers are required, it must be possible to draw high peak currents. Since ohmic losses increase as the square of the current but the light output increases only linearly, the average power limit probably will be lower for low-duty-cycle, high-peak-power operation than for continuous operation, although laser action is more readily achieved in pulsed operation.

The time resolution of conventional multiplier phototubes is indadeuate for the standard data rate PPM system, and is marginal at the reduced data rate (one-fifth standard). The crossed-field multiplier phototube has adequate time resolution but is probably slightly inferior in effective quantum efficiency because of the a-c bias employed. The traveling-wave phototube does not have adequate gain to avoid amplifier noise for very los signal levels. Since there are thus several devices which are almost adequate, a phototube of adequate time resolution and with noiseless amplification is hypothesized. Present-day devices are probably not deficient by more than a factor of two, and if the data rate is even further reduced the virtually ideal performance of the multiplier phototube may be fully realized. Unless time resolution is adequate to transmit pulses without smearing, peak pulse power is diminished and more signal power is required.

It should be noted that the bandwidth desired for PPM is approximately fourteen times that required for PCM; however, if bandwidth capability is not available for PPM, it is only necessary to transmit a slightly stronger signal, because the time of arrival of the leading edge of the pulse can be detected to within a small fraction of the pulse duration. It would be possible, and quite informative, to compare PCM and PPM for systems of equal bandwidth. In fact, the PCM system performance could theoretically be improved by transmitting the pulses in a short burst at the beginning of the sample period, although this is impractical for a continuous laser. The PPM system calculations were restricted to bandwidths such that the pulse duration equals the time resolution merely to limit the number of cases to be considered. Studies including different pulse durations and realistic pulse waveforms must be undertaken later to complete the understanding of this system.

From a system aspect the continuous laser PCM system is most nearly available. However, a factor of 1000 improvement in laser efficiency is desired; if the desired power output were to be obtained at present efficiency it would be too difficult to supply and remove the excess power. The hope of making such a technological advance makes this approach worth pursuing also.

5.0 SYSTEM PERFORMANCE CHARACTERISTICS

In this section those characteristics of the system which determine the power requirements of the transmitted beam are presented, and then this power is computed for several cases. For a mission to Mars, nighttime reception would be the rule, whereas for a mission to Venus daytime reception would be more common. The developments of lasers and detectors and the establishment of a particular mission will provide data which may be used with calculations of the type presented here to determine the optimum type of system and its characteristics.

5.1 Interfering Radiation

The major source of interfering radiation is the daytime sky. Measurements[10] of sky radiance from the visible region (0.5μ through 20μ) have been combined with the known short wavelength limit of solar radiation and converted to photon flux to provide a complete spectrum (Figure 19) of photon radiance. The near-infrared data represent the highest radiance among four spectra taken at 12,000 ft. elevation. The spectral radiance indicated at 0.63μ, converted to an elevation of about 2000 ft., gives a spectral photon radiance of 10^{16} photons $sec^{-1} cm^{-2} ster^{-1} \mu^{-1}$. From the known stellar magnitude and diameter of Mars at opposition, the corresponding figure for the surface of that planet is found to be 5×10^{15} photons $sec^{-1} cm^{2} ster^{-1} \mu^{-1}$.

5.2 Performance Calculations

The two items of chief concern are the d-c noise current due to the detection of sky light and the attenuation experienced by the laser beam in the transmission and detection process. The important parameters are N,

---

10. P. T. Vanderhei and B. J. Taylor, "Spectral Ground and Sky Backgrounds" and E. Bell, et al, "The Spectral Radiance of the Sky in the Colorado Springs Area," in Background Measurements During the Infrared Measuring Program 1956, ed. by M. R. Nagel, Geophysics Research Directorate, AF Combridge Research Laboratories, Air Force Research Division, USAF, Bedford, Massachusetts.

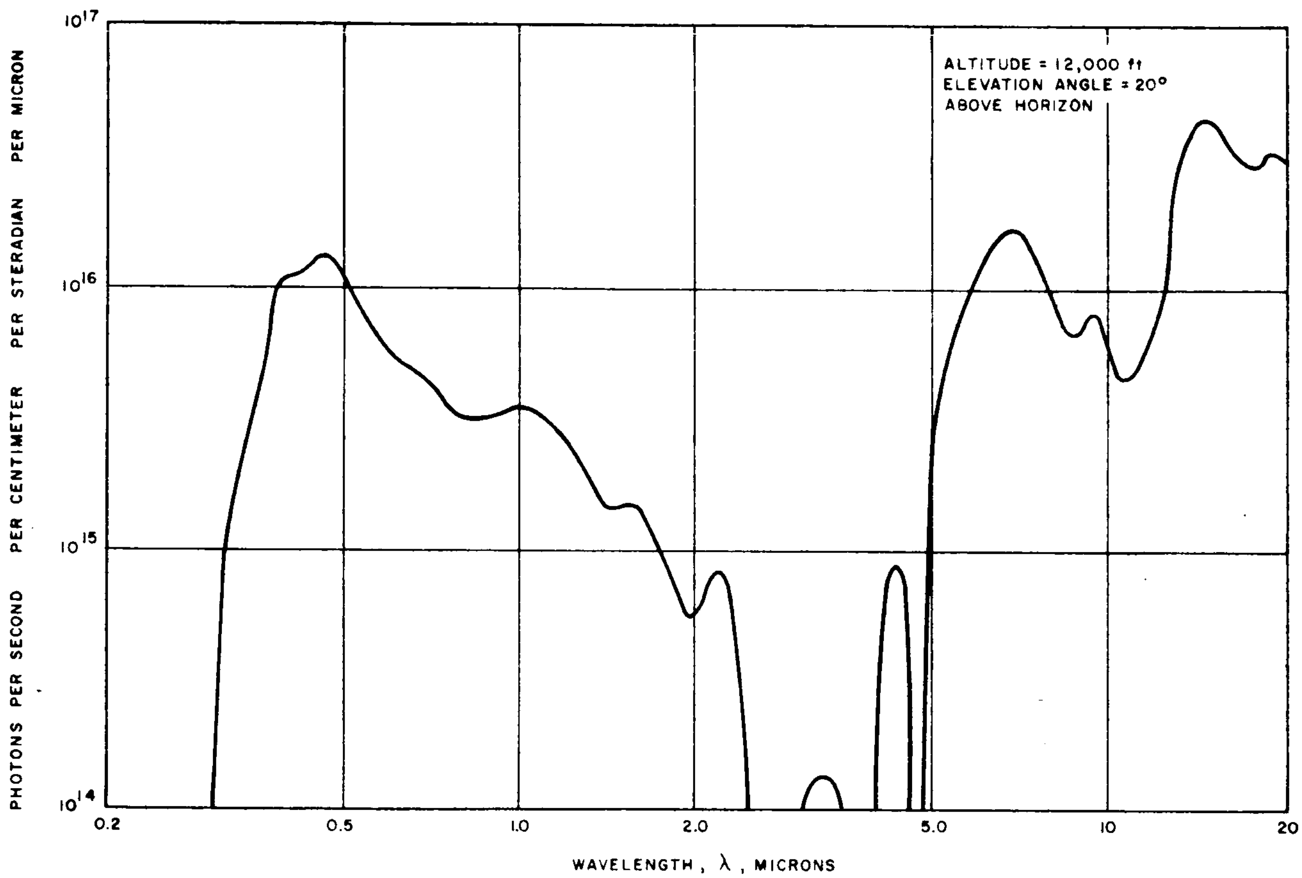


Figure 19. Spectral Photon Radiance of Daytime Blue Sky

the noise photoelectron count per sample period, and S_O, the signal photoelectron count per sample period. The key equations are

$$N = \frac{Q_B(\lambda)\Delta\lambda d_R^2 \theta_R^2 \pi^2 T_A T_R \eta}{16}$$

and

$$S_0 = \frac{d_R^2 T_T T_A T_R T_S \eta \tau}{\theta_T^2 R^2 (hc/\lambda)} P_T$$

where hc/λ is the energy of a photon, P_T is the average power output of the laser, and the other symbols are defined in Table II. The term T_A in the formula for N is omitted when the interference is due to sky light but when the interference is due to the radiation from a planet or star it must be included.

Several different values of N may be calculated, one for each appropriate combination of the different values of θ_R, τ, and $Q_B(\lambda)$ given in Table II. For each of the types of signal processing a value of S_0 may be obtained from Figures 11 and 13, and hence values of P_T may be determined. The results are presented in Table III. The design example is based on the reduced data rate and PPM. The required powers, 11.4 w, 7.4 w, and 4.0 w for the various conditions, can be obtained in this particular configuration only if laser efficiencies are approximately 0.53, 0.43 and 0.29 respectively because of the limited cooling capacity (10 w) hypothesized. If such efficiency cannot be obtained, the cooling system would have to be increased in capacity, and then it would probably be necessary to supply a heat-transfer liquid (e.g., oil) to remove the cooling system heat output from the telescope because radiative cooling would no longer be adequate. Further reduction in data rate, however, may reduce power transmiss on requirements and hence obviate the need for such additional cooling.

Table II. Parameter Values for the DVS-Earth Communication Link

Parameter	Symbol	Value
Distance	R	9.3×10^{12} cm
Transmitter beam diameter	θ_T	5×10^{-6} rad
Receiver field diameter	θ_R	
Daytime		10^{-4} rad
Nighttime		2.5×10^{-5} rad
Receiver aperture diameter	d_r	300 cm
Receiver optical filter noise bandwidth	$\Delta\lambda$	10^{-3} μ
Transmitter optical transmittance	T_T	0.85
Atmospheric optical transmittance	T_A	0.75
Receiver optical transmittance	T_R	0.60
Atmospheric scintillation factor	T_S	0.90
Receiver phototube quantum efficiency	η	0.05
Sample period	τ	
Standard data rate		10^{-7} sec
Reduced data rate		5×10^{-7} sec
Spectral photon radiance	$Q_B(\lambda)$	
Daytime sky		$10^{16} \frac{\text{photons}}{\text{sec cm}^2 \text{ sr} \mu}$
Planet Mars		$5 \times 10^{15} \frac{\text{photons}}{\text{sec cm}^2 \text{ sr} \mu}$
Nighttime Sky		$\cong 0 \frac{\text{photons}}{\text{sec cm}^2 \text{ sr} \mu}$
Signal photoelectrons per sample period	S_0	
Standard data rate		$0.227\ P_T$
Reduced data rate		$1.14\ P_T$
Error rate (approx., per sample)	R_E	10^{-2}

Table III. Photoelectron Counts and Required Transmitter Average Power

Circumstances	N	S_0	P_T (watts)
Reduced data rate			
Daytime sky	75		
PPM		13	11.4
Pol. PCM		98	85
Nighttime, Mars	2		
PPM		8.4	7.4
Pol. PCM		48	42
Nighttime sky	0		
PPM		4.6	4.0
Pol. PCM		41	36
Standard data rate			
Daytime sky	15		
PPM		10	44
Pol. PCM		64	2.92
Nighttime, Mars	0.4		
PPM		6.7	30
Pol. PCM		43	190
Nighttime sky	0		
PPM		4.6	20
Pol. PCM		41	180

Note that an average power of 11.4 w for the PPM system implies an average diode laser current of at least 5.7 a., and hence a peak current value of approximately 600 a. Such currents are not inconsistent with present diode laser practice, but do present bothersome problems in application to cooled devices and also in high-speed switching circuitry.

5.3 Summary of Theoretical Studies

Earlier in this study the question of the departure of laser signals from Poisson statistics was studied, and it was decided that probably no important departure from the classical Poisson statistics occurs for the low signal levels expected at detection. The work of Glauber[11] supports this contention. It is probably only in a special application such as the local oscillator for heterodyne detection, in which fluctuations as small as one part in 10^8 would be troublesome, that these excess fluctuations are observed. The dynamical fluctuations of the ordinary ruby laser are a different matter; these would probably have to be suppressed. It was therefore determined that the sort of analysis presented in this report is valid, and moreover that Jones' analysis[7] has greater applicability than he claimed for it.

It was next determined that many of the peculiar results presented in the literature were due to failure on the part of many authors to realize that attenuation of a beam of light is a random process occuring (ordinarily) one photon at a time and hence a Poisson distribution of photons (or photoelectrons) is always the result of greatly attenuating a steady beam of light. This result can be derived from general formulae[12] for amplification and attenuation of radiation. It is also readily understood as a partition noise[13]. No means of detection can avoid the noise due to this fluctuation. Partition creates phase fluctuations as well as amplitude fluctuations[14].

---

11. R. J. Glauber, "Photon Correlations," Phys. Rev. Lett. 10, 84-85, February 1, 1963.

12. K. Shimode, H. Takahasi, and C. H. Townes, "Fluctuations in Amplification of Quanta," J. Phys. Soc. Jap., 12, 686-700, June 1957.

13. H. A. Haus and J. A. Mullen, "Photon Noise," NEREM Record, November 6, 1962.

14. R. Serber and C. H. Townes, "Limits on Electromagnetic Amplification due to Complementarity," in Quantum Electronics, ed. by C. H. Townes (Columbia University Press, New York, 1960).

After establishing this firm basis for analysis, the study of specific systems was undertaken. It was found that the most favorable systems were those in which photon counts were small and hence Gaussian approximations were not adequate. In circumstances in which the dominant noise is that due to fluctuations in the signal itself, even the concept of signal-to-noise ratio is not very useful. Calculations based on direct determination of probabilities using tables of the Poisson distribution were found to be best; signal, noise, and threshold values were related by calculations similar to those employed by Jones[7], and two promising types of signal processing for optical transmission of information were evaluated.

5.4 Conclusions

It has been found that television can be transmitted in real time from a distance of 5×10^7 nautical miles by a hypothetical system employing components having performance not far beyond that which is available today. The system configuration depends upon the relative progress of various areas of technology. The study to date is not comprehensive in that not all possible variations of parameters have been analyzed; however, the general features of interplanetary optical communication systems of the future are represented in the designs presented in this report.